



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 177599

TO: Sheridan Swope
Art Unit: 1656
Location: rem/2B71/3C70
Serial Number: 09/940235

Friday, February 03, 2006

From: Beverly Shears
Location: Biotech-Chem Library
REM 1A54
Phone: 571-272-2528
beverly.shears@uspto.gov

Search Notes

Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

Published Applications Database - November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).



From: Swope, Sheridan
Sent: *CRFE* Thursday, January 26, 2006 3:36 PM
To: STIC-Biotech/ChemLib
Subject: RE: 09/940,235

PLEASE ALSO ALIGN SID 9, 10, 11, AND 12.
Thanks!

-----Original Message-----

From: Swope, Sheridan
Sent: Wednesday, January 25, 2006 8:28 PM
To: STIC-Biotech/ChemLib
Subject: 09/940,235

For 09/940,235, pls search:

SID 9 against the NT and AA databases *NA 1541*

SID 10 against the NT and AA databases *NA 1661*

SID 11 against the NT and AA databases *NA 1782*

SID 12 against the NT and AA databases *NA 2096*

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

IG
STN
Dialog
APS
Geninfo
SDC
DARC Questel
Other *CGW*

US-09-940-consensus	184	TCTCGATCCCGCGAAATTAATACGACTACTATATAGGAGACCAACACGGTTCTCCCTCTAGAG tctcgatccgcgaaattaatacgactcactataggagagaccacaacggtttccctctagag
US-09-940-consensus	1	TTTGTGTTAACTTTTAAGAAGGAGATATACCATGATaGCTGTGCTCCTGAaTGCTaCT
US-09-940-consensus	145	AATAATTTTGTGTTAACTTTTAAGAAGGAGATATACCATGATaGCTGACCTGAGTGCTGCT
US-09-940-consensus	195	AATAATTTTGTGTTAACTTTTAAGAAGGAGATATACCATGCTGCAAGCACACAGATTGTACC
US-09-940-consensus	245	AATAATTTTGTGTTAACTTTTAAGAAGGAGATATACCATGCTGCAAGCACACAGATTGTACC
US-09-940-consensus		ataaatTTTGTGTTAACTTTTAAGAAGGAGATATACCATG-Tg--G-aC--Ag-----TaC-
US-09-940-consensus	56	AGA-CGTCCCTCTGTaAAaAACAGCCAAATTGGTTGTAGCGTTGCTGTGACTGTTGAGGGG
US-09-940-consensus	206	AGAcCGTCCaTCTGTcAAcAACAGCCAAATTGGTTGTAGCGTTGCTGTGACTGTTGAGGGG
US-09-940-consensus	256	CATAGCTGAGAAGTGTTTTCATCATGCTCTGTGGACTTCCTATGTGTCGGAGAAACGTGG
US-09-940-consensus	306	CATAGCTGAGAAAGTGTTTTCATCATGCTCTGTGGACTTCCTATGTGTCGGAGAAACGTGG
US-09-940-consensus		---a-T--g-----t-t-t-A-----G-----T--C--TG-G-----G-----G-G
US-09-940-consensus	117	ACGAATCAAGACATTAGTCTTAAATTTTGTGAATCGATTAACATCAGACCTGCTCATG
US-09-940-consensus	267	ACGAATCAAGACATTAGTCTTAAATTTTGTGAATCGATTAACATCAGACCTGCTCATG
US-09-940-consensus	317	GAGAAGCGACGGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAG
US-09-940-consensus	367	GAGAAGCGACGGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAG
US-09-940-consensus		--GAA---AG-----TT--A-----CA-C-----C-CA-G
US-09-940-consensus	178	GAGGAAAGACAGAGCAGGCTTAAGTCCAAATCAAAACCATTTGCTACTGATAGTGGCGC
US-09-940-consensus	328	GAGGAAAGACAGAGCAGGCTTAAGTCCAAATCAAAACCATTTGCTACTGATAGTGGCGC
US-09-940-consensus	378	GACATCCTATAGAAATTTGGAGACACCTTGGAGCAGAGAGATATCGAGGAAACCTGCTCCAG
US-09-940-consensus	428	GACATCCTATAGAAATTTGGAGACACCTTGGAGCAGAGAGATATCGAGGAAACCTGCTCCAG
US-09-940-consensus		GA-----A-AGA---G-----A-T--A--A-AA-----A-T-G-----G--C--
US-09-940-consensus	239	GATGTCAATAAACTTGAGAAAGCTCACTTACTTAAGAGCTATTTCAGAACAAATTGATCGCT
US-09-940-consensus	389	GATGTCAATAAACTTGAGAAAGCTCACTTACTTAAGAGCTATTTCAGAACAAATTGATCGCT
US-09-940-consensus	439	TGCATCTTGCAAGGCAACCGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGA
US-09-940-consensus	489	TGCATCTTGCAAGGCAACCGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGA
US-09-940-consensus		----TC-----A-----G---G---GA-----AA--G--A-----AC---TG--C--
US-09-940-consensus	300	AAAGTCCACAGTAACGACGACTACTTTGAGGTCAATTGATTTTGGCAAGCGATGCAACCAATTA
US-09-940-consensus	450	AAAGTCCACAGTAACGACGACTACTTTGAGGTCAATTGATTTTGGCAAGCGATGCAACCAATTA
US-09-940-consensus	500	CCACATCGACGCGATCTGGCCCTTCAACCGATGTTCTGATATGCTGGACCTGAGTGGCTGCT
US-09-940-consensus	550	CCACATCGACGCGATCTGGCCCTTCAACCGATGTTCTGATATGCTGGACCTGAGTGGCTGCT
US-09-940-consensus		-----C-----A-----G-C---T-----G-T--T-G-----C-----C-----
US-09-940-consensus	361	CTGATCGAAACGGCAAGGTCTACTTTGCTGCAACAGATGTTTCGGTTAACTTGGCGACCCA

US-09-940- 511 CTGATCGAAGCGGCAAGGTCTACTTTGCTGACAAAGANGGTTGCGTAACTTGGCCGACCCA
US-09-940- 561 AGACCGTCCATCTGTCAACCAACAGCAAAATGGTTGTAGCGTTGCTGGTACTGTTGAGGGG
US-09-940- 611 AGACCGTCCATCTGTCAACCAACAGCAAAATGGTTGTAGCGTTGCTGGTACTGTTGAGGGG
consensus -----A-----T-----G--T--T-----TG--GA-----
US-09-940- 422 ACCTGTCCAAAGAAATTTTGCTAAGCGGCATCTGCGGTTAGACCATATAAGAAAAACCA
US-09-940- 572 ACCTGTCCAAAGAAATTTTGCTAAGCGGCATCTGCGGTTAGACCATATAAGAAAAACCA
US-09-940- 622 ACGAATCAAGACATTTAGTCTTAAATTTTGAATCGATCAACATCAAGCCTGCTCATG
US-09-940- 672 ACGAATCAAGACATTTAGTCTTAAATTTTGAATCGATCAACATCAAGCCTGCTCATG
consensus AC---TC-A---ATT--T--TAA-----T-----AC---A--A-----
US-09-940- 483 ATACAAAACCAAGCGAAATCTGTGATGGGAATATCTGTACAGTTTACTCCCTTAAACC
US-09-940- 633 ATACAAAACCAAGCGAAATCTGTGATGGGAATATCTGTACAGTTTACTCCCTTAAACC
US-09-940- 683 GAGGAACACAGACAGCAAGGCTTAAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGGC
US-09-940- 733 GAGGAACACAGACAGCAAGGCTTAAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGGC
consensus ----AAA--CA-----AA--CT--G-----A--A--A-----G-T-----T-----C
US-09-940- 544 CTGATGACGATTTAGACAGCGCTCTCAAGACTATAAGCTATTGAAAACTAGCTATCGG
US-09-940- 694 CTGATGACGATTTAGACAGCGCTCTCAAGACTATAAGCTATTGAAAACTAGCTATCGG
US-09-940- 744 GATGTCAATAAACTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCT
US-09-940- 794 GATGTCAATAAACTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCT
consensus -----T-AC-----C-----A-----CT-A--A--A--A--GCTATT-AA-A-----
US-09-940- 605 TGACACCATCATCTCAAGAAATTTACTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAAC
US-09-940- 755 TGACACCATCATCTCAAGAAATTTACTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAAC
US-09-940- 805 AACGTCCACAGTAACGACGACTACTTTGAGGTCATTTGCAAGCGATGCAACCATTA
US-09-940- 855 AACGTCCACAGTAACGACGACTACTTTGAGGTCATTTGCAAGCGATGCAACCATTA
consensus -----CCA-----A-C-----T-----A-----A-----A--AA-----
US-09-940- 666 CCAGGTATACGATTTTGAACGCTGACTCTCAATCGTCACTCATGACAAATGACATTTCC
US-09-940- 816 CCAGGTATACGATTTTGAACGCTGACTCTCAATCGTCACTCATGACAAATGACATTTCC
US-09-940- 866 CTGATCGAAACGGCAAGGTCTACTTTGCTGACAAAGATGGTTGCGTAACTTTGCCGACCCA
US-09-940- 916 CTGATCGAAACGGCAAGGTCTACTTTGCTGACAAAGATGGTTGCGTAACTTTGCCGACCCA
consensus C-----C-A-ACG-----T-----CAA---T-----TC---A---T---C-----C-----
US-09-940- 727 GTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTA
US-09-940- 877 GTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTA
US-09-940- 927 ACCTGTCCAAGAAATTTTGTCTAAGCGGCATCTGCGCGTTAGACCATATAAGAAAAACCA
US-09-940- 977 ACCTGTCCAAGAAATTTTGTCTAAGCGGCATCTGCGCGTTAGACCATATAAGAAAAACCA
consensus -----G-----T-----NAG-G---A--T-C-----T--A--AT-----A--AA-----A

US-09-940- 788 TAGGATCAATAAAAAATCTGCTCTGAATGAAGAAATAAACAAACACTGACCTGCTCTCTGAG
US-09-940- 938 TAGGATCAATAAAAAATCTGCTCTGAATGAAGAAATAAACAAACACTGACCTGCTCTGAG
US-09-940- 988 ATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACC
US-09-940- 1038 ATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACC
consensus -----A--A--AA---A--T---TGA-----A-ATA-----ACA-T---CT---T-----
US-09-940- 849 AAATATTACGTCCTTTAAAAAGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA
US-09-940- 999 AAATATTACGTCCTTTAAAAAGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA
US-09-940- 1049 CTGATGACGATTTTCAAGCCAGGCTCTCAAGATATCTAAAGTATTTGAAAAACACTAGCTATCGG
US-09-940- 1099 CTGATGACGATTTTCAAGCCAGGCTCTCAAGATATCTAAAGTATTTGAAAAACACTAGCTATCGG
consensus -----T-----A--A--G---AAG-----T-----A--C-C--C--T--G--G--
US-09-940- 910 AACTGTTACCATCAAAATAGTTGATGTCGATACCAAGAAATGCTTAAAAAGTGAGCAGCT
US-09-940- 1060 AACTGTTACCATCAAAATAGTTGATGTCGATACCAAGAAATGCTTAAAAAGTGAGCAGCT
US-09-940- 1110 TGACACCATCACTCAAGAAATTTACTAGTCAAGCAACAAGCAATTTTAAACAAAAACCAAC
US-09-940- 1160 TGACACCATCACTCAAGAAATTTACTAGTCAAGCAACAAGCAATTTTAAACAAAAACCAAC
consensus -----C-----A---T-A-----A---AC-AA-----T---AAA-----A-----
US-09-940- 971 CTTAAACAGCTAGCGAAACGTAACTTAGACTTTCAGAGATTTATACGATCCTCGTGATAAGGCT
US-09-940- 1121 CTTAAACAGCTAGCGAAACGTAACTTAGACTTTCAGAGATTTATACGATCCTCGTGATAAGGCT
US-09-940- 1171 CAGGCTATACGATTTTATGAACGCTGACTCTCAATCGTCACTCATGCAATGACATTTTCC
US-09-940- 1221 CAGGCTATACGATTTTATGAACGCTGACTCTCAATCGTCACTCATGCAATGACATTTTCC
consensus C-----C-----G-----NAC-T-----A---T-A--C-----C-----C-----
US-09-940- 1032 AAATCTACTTCAACCAATCTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAG
US-09-940- 1182 AAATCTACTTCAACCAATCTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAG
US-09-940- 1232 GTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTAAATAATCGGGAACAAGCTTTA
US-09-940- 1282 GTACGATTTTACCAATGGATCAAGATTTTACTTACCGTGTAAATAATCGGGAACAAGCTTTA
consensus --AC-A-T-TAC-A-----TC-A---TTT---TA---TG-----A-A-C-----A-----
US-09-940- 1093 TAGAGGATAATACGATGACCAACCGTATCATAAACCGTTTATATGCGGAAGCGACCCGA
US-09-940- 1243 TAGAGGATAATACGATGACCAACCGTATCATAAACCGTTTATATGCGGAAGCGACCCGA
US-09-940- 1293 TAGAGTCAATAAAAAATCTGCTCTGAATGAAGAAATAAACCAACACTGACCTGATCTCTGAG
US-09-940- 1343 TAGAGTCAATAAAAAATCTGCTCTGAATGAAGAAATAAACCAACACTGACCTGATCTCTGAG
consensus TAG-----A---A--AT-----C-----A--A-----TG-C-----C-----
US-09-940- 1154 AGGAGAGAAATGCTAGCTACCATTTAGCTGTGGTGGC-CAGCGCAACAGATGTA Ccata
US-09-940- 1304 AGGAGAGAAATGCTAGCTATCATTTAGCGCGTGGTGGT-CAGCGCAGCAAAATGgt Cagccc
US-09-940- 1354 AAATATTACGTCCTTTAAAAAGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA

```
US-09-940- 1404 AATATTATGTCCTTAAATAAGGGAAAGCCGTATGATCCCTTTGATCGCATCACTTGA
consensus  A-----A---C-----Aa-A-----A-G--GT--g-----C---g--c--aGTC-cttga

US-09-940- 1215 gctgagaagtgtttTGatcatgctgctggGaCctccTaTgtGtCggaGaaAcgtgggAga
US-09-940- 1365 cAGTccccCGgtggCgtcAGTcCaAaAgcaAGcCcggTgTtATGaCaAATGgAAaacaCtAtc
US-09-940- 1415 AACTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTGCTTAAAAAAGTGACAGCT
US-09-940- 1465 AACTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTGCTTAAAAAAGTGACAGCT
consensus  aactgttc-ccatc--atac-ttgatgtcGa-acca-c-aat-gCtaa-aA-tgagca-ct

US-09-940- 1276 AGccctAcCAAggcTGGatGatGgtagAttgacttgcTggaGaaGcAgcgGacgcat
US-09-940- 1426 AGatAaAtCAAcagTGGAGcGgGacctAcctAGgTAAtgTgtGGTttGtActtGttatgg
US-09-940- 1476 CTTAAcAGCTAGCGAAcGTAACTTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCT
US-09-940- 1526 CTTTACAGCTAGCGAAcGTAACTTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCT
consensus  T-taacAgC-Agcg---gt-ac-taga-ttcaga-attt--acg-tcc-c-tga-aaggct

US-09-940- 1337 caecttGcacttcTagaAtaGatGcAaGAtCagGNaCaaggacatCctatagaattgga
US-09-940- 1487 AggaagCcggagTtttAACTGCGaaAGTAAaCctGAAGCtgaagagacTTgctTtGAcAAG
US-09-940- 1537 AAACTACTCTCAACAACACTCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAaAG
US-09-940- 1587 AAACACTCTCAACAACACTCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAaAG
consensus  aaact-Ctctac-acaaA-ct-gatg-ttt-g-ta--at-gactatacCtTaaactggaaaaag

US-09-940- 1398 GACACctGGAGCAagaAGgataaatcGaGaaAacctgctcCagtGcAtctGCaCaggcAAcG
US-09-940- 1548 TACACtggGAAcACtTAcCgAgTggGtGacACtTAtgaCGTcctAaaGaCtCCatGAtCt
US-09-940- 1598 TAGAGGATAAATCACGATGACACCAACCGTATCATAAcCGTTTATATGGCAAGCGACCCGA
US-09-940- 1648 TAGAGGATAAATCACGATGACACCAACCGTATCATAAcCGTTTATATGGCAAGCGACCCGA
consensus  tA-A-ga--AtCAcga-gacaccaa-cGta-cataaccg-ttata-gggc-a-cgac-c-a

US-09-940- 1459 GccgagGagagTgGAagtGtGagaGGCaacacctctgTgcagaccACatcgagcggaatctgg
US-09-940- 1609 GGGActGtAccTgCAtcgGgcTgGGcGagggagAaTAgctGtACcaTcTAA
US-09-940- 1659 AGGAGAGAAATGCTAGCTAtCATTTAGC ctAtgAtaaGatCgtTATAccgaagAag
US-09-940- 1709 AGGAGAGAAATGCTAGCTACCAATTTAGCtggTggccaggcGcaacagATtgtacccAta
consensus  -ggagaGaagt--a-ctagc-ttt-GC--g-tagtaa-gagcacc-tatac-ga-ca-g

US-09-940- 1520 ccccttcaccgatgttcgtagg
US-09-940- 1662
US-09-940- 1715 aacGAGAAgT
US-09-940- 1770 gctGAGAAgTgttttgatcatgctgctgggaacttccctatgtggtcgggaaaaacgtgggaga
consensus  -ccgagaagtg-t-t-----gctgctgggaacttccctatgtggtcgggaaaaacgtgggaga

US-09-940- 1542
```

```
US-09-940- 1662
US-09-940- 1725 ttaCAGC
US-09-940- 1831 agccctaccaaggctggatgtagattgtacttgcctggggaaggCAGCggaagcat
consensus
US-09-940- 1542
US-09-940- 1662
US-09-940- 1732
US-09-940- 1892 cacttgcacttctagaaatagatgcaacgatcaggacacacaaggacatcctatagaattgga
consensus cacttgcacttctagaaatagatgcaacgatcaggacacacaaggacatcctatagaattgga
US-09-940- 1542
US-09-940- 1662
US-09-940- 1732 tACCTGC gtTatACAGG
US-09-940- 1953 gacacctggagcaagaaggataatcgaggaaACCTGctccagtgcatcTgcACAGGcaacg
consensus gacacctggagcaagaaggataatcgaggaa-acctgctccagtgca--t--acaggcaacg
US-09-940- 1542
US-09-940- 1662
US-09-940- 1749 GACACCT ataccTgata
US-09-940- 2014 gccgaggagagtggaaagtgtgagaggcACACTctgtgcagaccacacatcgagcgaTctgg
consensus gccgaggagagtggaaagtgtgagagg-acacctctgtgcagaccacacatcga-----t-----

US-09-940- 1542
US-09-940- 1662
US-09-940- 1766 aCCCTaacgacAaaTaa
US-09-940- 2075 cCCCTtcaccgAtgTtcgtag
consensus -ccct-----a--t--gttag
```

Alignment score = -2900.00

Scoring matrix:

6	7	8	9
6	287	-1173	-2369
7		-423	-830
8			146
9			

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:53:18 ; Search time 7834.07 Seconds
(without alignments)
11181.386 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 1541
Sequence: 1 ttgtttaacttaagaagg.....ccttcaccgatgttcgtag 1541

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sta.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1134	73.6	2385	6	AR143998 Sequence
2	1132.6	73.5	1242	6	AR144000 Sequence
3	1132.6	73.5	1401	6	I05204 Sequence 5
4	1132.6	73.5	2568	1	STRSKC
5	1132.6	73.5	2568	6	A04926
6	1132.6	73.5	8931	1	SEDEXB
7	1127.8	73.2	2568	6	E00522
8	1120.2	72.7	7057	6	CQ979820
9	1118.6	72.6	2566	6	AR068768
10	1111.8	72.1	1245	1	S46536
11	1111.6	72.1	1257	6	A20015
12	1111.6	72.1	1257	6	I13203
13	1110.2	72.0	1407	6	E01413
14	1110.2	72.0	1512	6	A20016
15	1110.2	72.0	1512	6	I13204
16	1108.6	71.9	1317	6	A20009
17	1108.6	71.9	1317	6	I13197
18	1108.6	71.9	1335	6	A20006

19	1108.6	71.9	1335	6	I13194	I13194 Sequence 14
20	1108.6	71.9	1458	6	A20027	A20027 SEQ ID NO:
21	1108.6	71.9	1458	6	I13215	I13215 Sequence 42
22	1108.6	71.9	1467	6	A20030	A20030 SEQ ID NO:
23	1108.6	71.9	1467	6	I13218	I13218 Sequence 46
24	1108.6	71.9	2588	6	A20021	A20021 SEQ ID NO:
25	1108.6	71.9	2589	6	I13209	I13209 Sequence 34
26	1100.6	71.4	1473	1	SGSKG	XI3400 Streptococc
27	1099	71.3	1311	1	AY368335	AY368335 Streptoco
28	1084.4	70.4	1209	6	AR175891	AR175891 Sequence
29	1084.4	70.4	1209	6	AX030315	AX030315 Sequence
30	1084.4	70.4	1245	6	AR175892	AR175892 Sequence
31	1084.4	70.4	1245	6	AX030316	AX030316 Sequence
32	1083.8	70.3	1122	6	AR175893	AR175893 Sequence
33	1083.8	70.3	1122	6	AX030317	AX030317 Sequence
34	1083.8	70.3	1158	6	AR175894	AR175894 Sequence
35	1083.8	70.3	1158	6	AX030318	AX030318 Sequence
36	1078.8	70.0	2253	6	I13213	I13213 Sequence 39
37	1077.6	69.9	1119	6	I13206	I13206 Sequence 30
38	1075.8	69.8	2252	6	A20025	A20025 SEQ ID NO:
39	1065.6	69.1	1118	6	A20018	A20018 SEQ ID NO:
40	1066.2	65.3	1323	1	AY234129	AY234129 Streptoco
41	1066.2	65.3	1323	1	AY234137	AY234137 Streptoco
42	1066.2	65.3	50354	1	AE014169	AE014169 Streptoco
43	1066.2	65.3	110000	1	BA000034_16	Continuation (17 o
44	1066.2	65.3	110000	1	BA000034_17	Continuation (18 o
45	999.8	64.9	1323	1	AY234136	AY234136 Streptoco

ALIGNMENTS

RESULT 1
LOCUS AR143998 2385 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6210667.
ACCESSION AR143998
VERSION AR143998.1 GI:15105865
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2385)
AUTHORS Reed G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 1 03-APR-2001;
FEATURES Location/Qualifiers
source 1..2385
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	73.6%	Score 1134;	DB 6;	Length 2385;
Best Local Similarity	98.7%	Pred. No. 1.4e-278;		
Matches 1143;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;
Qy	30	ATGATAGCTGGTCTCTGAATGGCTACTAGATCGTCTTCTGTGTAATAACAGCAATTGGTT	89	
Db	1141	AGGATTCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTAGTT	1200	
Qy	90	GTTAGCGTTCGTGCTACTGTTAGGGGCGCAATCAAGACATTAGCTTTAAATTTTGA	149	
Db	1201	GTTAGCGTTCGTGCTACTGTTAGGGGCGCAATCAAGACATTAGCTTTAAATTTTGA	1260	
Qy	150	ATCGATCTAAACATCACCACTGCTCTATGGAGGAAGACAGCAAGCGCTTAAGTCCAAA	209	
Db	1261	ATCGATCTAAACATCACCACTGCTCTATGGAGGAAGACAGCAAGCGCTTAAGTCCAAA	1320	
Qy	210	TCAAAACCAATTTGCTACTGATAGTGGCGCGATGTGCACATAAATTGAGAAGCTGACTTA	269	
Db	1321	TCAAAACCAATTTGCTACTGATAGTGGCGCGATGTGCACATAAATTGAGAAGCTGACTTA	1380	
Qy	270	CTAAAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTACAGGACTACTTTTGAG	329	

```
Db 1381 CTAAGAGGCTATTCAAGAAACAATTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAG 1440
Qy 330 GTCAATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACCGCAAGGCTCTACTTTGCT 389
Db 1441 GTCAATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACCGCAAGGCTCTACTTTGCT 1500
Qy 390 GACAAAGATGGTTCGGTAACTTTCGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGA 449
Db 1501 GACAAAGATGGTTCGGTAACTTTCGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGA 1560
Qy 450 CATGTGCGGTTAGACCATATAAAGAAACCAATACAAACCAACGAACTGTTGAT 509
Db 1561 CATGTGCGGTTAGACCATATAAAGAAACCAATACAAACCAACGAACTGTTGAT 1620
Qy 510 GTGGATATACGTACAGTTCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTC 569
Db 1621 GTGGATATACGTACAGTTCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTC 1680
Qy 570 AAAGATATACGTATTTGAAACCACTAGCTATTCGGTGAACACCATCAATCTCAAGAAATTA 629
Db 1681 AAAGATATACGTATTTGAAACCACTAGCTATTCGGTGAACACCATCAATCTCAAGAAATTA 1740
Qy 630 CTAGCTCAAGCAACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 689
Db 1741 CTAGCTCAAGCAACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 1800
Qy 690 GACTCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGATCAAA 749
Db 1801 GACTCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGATCAAA 1860
Qy 750 GAGTTTACTTACCGTGTAAACATCGGGAACAGCTTTATAGGATCAATAAAATCTGGT 809
Db 1861 GAGTTTACTTACCGTGTAAACATCGGGAACAGCTTTATAGGATCAATAAAATCTGGT 1920
Qy 810 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAA 869
Db 1921 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAA 1980
Qy 870 GGGGAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCAATCAATAC 929
Db 1981 GGGGAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCAATCAATAC 2040
Qy 930 GTTGATGTGATACCAACCAATTTGCTTAAAGTGACAGCTCTTACAGCTAGCGAACGT 989
Db 2041 GTTGATGTGATACCAACCAATTTGCTTAAAGTGACAGCTCTTACAGCTAGCGAACGT 2100
Qy 990 AACTTAGACTTCAGAGATTTATACGATCCTCGTGTAAAGCTAAACTTACTCTCAACAAT 1049
Db 2101 AACTTAGACTTCAGAGATTTATACGATCCTCGTGTAAAGCTAAACTTACTCTCAACAAT 2160
Qy 1050 CTCGATGCTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGAT 1109
Db 2161 CTCGATGCTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGAT 2220
Qy 1110 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGAAGCGAGAGATGCTAGC 1169
Db 2221 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGAAGCGAGAGATGCTAGC 2280
Qy 1170 TACCATTAGCTGGTGGT 1187
Db 2281 TATCAATTTAGCCTATGAT 2298
```

```
RESULT 2
LOCUS AR144000 AR144000 1242 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6210667.
ACCESSION AR144000
VERSION AR144000.1 GI:15105867
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
```

```
Unclassified.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Reed,G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 5 03-APR-2001;
FEATURES Location/Qualifiers
source 1..1242
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 73.5%; Score 1132.6; DB 6; Length 1242;
Best Local Similarity 98.8%; Pred. No. 3.1e-278;
Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 ATAGCTGGTCTCTGAAATGGCTACTAGATCGTCTTCTGTAAATTAACGCCAATTTGGTTGTT 92
Db 1 ATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTAGTTGTT 60
Qy 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
Qy 153 GATCTAACATCAACGACCTGCTCATGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 212
Db 121 GATCTAACATCAACGACCTGCTCATGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 180
Qy 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 272
Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 240
Qy 273 AAGGCTATTTCAAGAAACAATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGTC 332
Db 241 AAGGCTATTTCAAGAAACAATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGTC 300
Qy 333 ATTGATTTTTCGAAGCGATGCAACCATTTACTGATGCAAAACGGCAAGGCTCTACTTTGCTGAC 392
Db 301 ATTGATTTTTCGAAGCGATGCAACCATTTACTGATGCAAAACGGCAAGGCTCTACTTTGCTGAC 360
Qy 393 AAGATGTTTTCGGTAACTTTGCGGACCCAACTGTCCAGAAATTTTTCGTAAGCGGACAT 452
Db 361 AAGATGTTTTCGGTAACTTTGCGGACCCAACTGTCCAGAAATTTTTCGTAAGCGGACAT 420
Qy 453 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 512
Db 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 480
Qy 513 GAATATACGTACAGTCTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACAGGCTCTCAA 572
Db 481 GAATATACGTACAGTCTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACAGGCTCTCAA 540
Qy 573 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 632
Db 541 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 600
Qy 633 GCTCAAGCACAAAGCAATTTTAAACAAACCAACCGGCTATAGGATTTATGAACGTCAC 692
Db 601 GCTCAAGCACAAAGCAATTTTAAACAAACCAACCGGCTATAGGATTTATGAACGTCAC 660
Qy 693 TCCTCAATCGTCACTCATGACATGACATTTTTCGGTACGATTTTACCCTAAGGATCAAGAG 752
Db 661 TCCTCAATCGTCACTCATGACATGACATTTTTCGGTACGATTTTACCCTAAGGATCAAGAG 720
Qy 753 TTTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 721 TTTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
Qy 813 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 872
Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 840
Qy 873 GAAAGCCGTATCATCTCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 932
```

```
Db      841  GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTCCACCATCAATATACGTT 900
Qy      933  GATGTGATACCAACGAATTTGCTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 992
Db      901  GATGTGATACCAACGAATTTGCTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 960
Qy      993  TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAACAATCTC 1052
Db      961  TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAACAATCTC 1020
Qy      1053  GATGCTTTTGGTATTTATGACTACTACCTTAACCTGGAAGATGAGGATATACGATGAC 1112
Db      1021  GATGCTTTTGGTATTTATGACTACTACCTTAACCTGGAAGATGAGGATATACGATGAC 1080
Qy      1113  ACCAACCGTATCATACCGTTTATATGCGGCAAGCGACCCGAAGGAGAGATGCTAGCTAC 1172
Db      1081  ACCAACCGTATCATACCGTTTATATGCGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1140
Qy      1173  CATTTAGCTGGTGGT 1187
Db      1141  CATTTAGCTATGAT 1155

RESULT 3
I05204
LOCUS      105204
DEFINITION Sequence 5 from Patent EP 0248227.
ACCESSION I05204
VERSION I05204.1 GI:591209
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1401)
AUTHORS Hagenson,M.J. and Stroman,D.W.
TITLE Yeast production of streptokinase
JOURNAL Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES
    source
        1. .1401
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 73.5%; Score 1132.6; DB 6; Length 1401;
Best Local Similarity 98.4%; Pred. No. 3.1e-278;
Matches 1144; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy      25  ATACCATGATAGCTGCTCGTGAATGGCTACTAGATCGTCTCTCTGTAATAAACAAGCCAAAT 84
Db      2  ATTCCATGATTTGCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAAT 61
Qy      85  TGGTTGTTAGCTGTTGCTGTTAGTGGGACGAATCAAGACATTTAGTCTTAAATTTT 144
Db      62  TAGTTGTTAGCTGTTGCTGTTAGTGGGACGAATCAAGACATTTAGTCTTAAATTTT 121
Qy      145  TTGAAATCCGATCTAATCAGACCTGCTCATGAGGAGGAGACAGACGACGAGCTTAAGTC 204
Db      122  TTGAAATCCGATCTAATCAGACCTGCTCATGAGGAGGAGACAGACGACGAGCTTAAGTC 181
Qy      205  CAAATCAAAACCATTTGCTGATGATGAGTGGCGGATGTCACATAAACTTGAGAAAGCTG 264
Db      182  CAAATCAAAACCATTTGCTGATGATGAGTGGCGGATGTCACATAAACTTGAGAAAGCTG 241
Qy      265  ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACGTCCTCAAGTAAACGAGCTACT 324
Db      242  ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACGTCCTCAAGTAAACGAGCTACT 301
Qy      325  TTGAGGTCAATTGATTGCAAGGATGCAACCTTACTGATCGAAACGGCAAGGCTTACT 384
Db      302  TTGAGGTCAATTGATTGCAAGGATGCAACCTTACTGATCGAAACGGCAAGGCTTACT 361
Qy      385  TTGCTGACAAAGATGTTGCTGATTAACCTTGGCGACCAACCTGTCGAAGATTTTCTTAA 444
```

```
Db      362  TTGCTGACAAAGATGGTTCCGTTAACCTTGGCCGACCCCAACCTGTCCAGAAATTTTTGCTAA 421
Qy      445  GCGGACATGTGCGGCTTAGACCATATAAGAAAAACCAATACAAACCAACGGAATCTG 504
Db      422  GCGGACATGTGCGGCTTAGACCATATAAGAAAAACCAATACAAACCAACGGAATCTG 481
Qy      505  TTGATGTGAATATACCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAG 564
Db      482  TTGATGTGAATATACCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAG 541
Qy      565  GTCTCAAGATACCTAAGCTATTGAAACACATAGCTATCGGTGACACCAATCATCTCAAG 624
Db      542  GTCTCAAGATACCTAAGCTATTGAAACACATAGCTATCGGTGACACCAATCATCTCAAG 601
Qy      625  AATTACTAGCTCAAGACCAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATG 684
Db      602  AATTACTAGCTCAAGACCAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATG 661
Qy      685  AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 744
Db      662  AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 721
Qy      745  ATCAAGATTTTACTTACCGTGTAAATAATCGGGAACCAAGCTTATAGGATCAATAAAAAT 804
Db      722  ATCAAGATTTTACTTACCGTGTAAATAATCGGGAACCAAGCTTATAGGATCAATAAAAAT 781
Qy      805  CTGCTCTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 864
Db      782  CTGCTCTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 841
Qy      865  AAAAAAGGGAAGCCGTATGATCCCTTTGATCGCAGTCACCTGAAACTGTTTACCACATCA 924
Db      842  AAAAAAGGGAAGCCGTATGATCCCTTTGATCGCAGTCACCTGAAACTGTTTACCACATCA 901
Qy      925  AATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCG 984
Db      902  AATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCG 961
Qy      985  AACGTAACCTTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTAAACTACTCTACA 1044
Db      962  AACGTAACCTTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTAAACTACTCTACA 1021
Qy      1045  ACAATCTCGATGCTTTTGGTATTATGACTATATCTTAACTTGGAAAAAGTAGAGGATATC 1104
Db      1022  ACAATCTCGATGCTTTTGGTATTATGACTATATCTTAACTTGGAAAAAGTAGAGGATATC 1081
Qy      1105  ACGATGACACCAACCGCTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGATG 1164
Db      1082  ACGATGACACCAACCGCTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGATG 1141
Qy      1165  CTAGCTACCAATTTAGCTGGTGGT 1187
Db      1142  TCAGCTATCATTTAGCTATGAT 1164
```

RESULT 4

```
STRSKC
LOCUS      2568 bp DNA linear BCT 26-APR-1993
DEFINITION Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
ACCESSION K02986
VERSION K02986.1 GI:153808
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
REFERENCE 1 (bases 1 to 2568)
AUTHORS Malke,H., Roe,B. and Ferretti,J.J.
TITLE Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A
JOURNAL Gene 34 (2-3), 357-362 (1985)
PUBMED 2989113
COMMENT Original source text: S.equisimilis (strain H46A) DNA, clone pMF5.
```

Draft entry and hard copy of sequence for [1] kindly provided by J.J.Ferretti, 03-SEP-1985.
The -35 and -10 regions are located at positions 760-765 and 781-786 respectively and an SD sequence at 808-813. Downstream from the coding region inverted repeats (positions 2176-2190 and 2203-2217) are thought to function as transcription terminators. The nucleotide sequence of skc does not support the hypothesis that the gene has evolved by duplication and fusion, as suggested by internal two-fold AA homologies of its product.

FEATURES

Location/Qualifiers
1. .2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
794_ .>2141
/product="skc mRNA"
819_ .2141
/note="prestreptokinase"
/codon_start=1
/transl_table=11
/protein_id="AAA26974.1"
/db_xref="GI:153809"
/translation="MKNYLSFGMFALLFALFTGVNSVOAIGPEWLLDRPSVNNSQL
VYSVAGTVGTQNDLSLKEFDLTSRPAHGGKTEQGLSPKSPFADSGAMSHKLEK
ADLLKAIQRLIANVHSDNDYFEVIDFASDAITDRNGKVFADKGSVLTLPQVQE
FLLSGHVRVPYKEKPIONAKSVDEYVTFPLNPDDEFRPLKDTLLKLTALGD
TITSOELLAAQOSILKNHGYTIYERDSSIYTHDNDIERTILPMDQBYRYRVNREQ
AYRNKKSGLNEINNDLISEKYIVLKGEKPYDFPDRSHLKLFTIKYVDVDNELL
KSEQLLTASERINLDFDLDPDKLNNLDLAFGIMDYTLTGKVEDNHDNDNRIT
VYMKRPEGENASYHLAYDKRYTEBEREYVSYLRYTGPDPNPDK"
819_ .896
/note="streptokinase signal peptide"
mat_peptide
897_ .2138
/product="streptokinase"
ORIGIN 5 bp upstream of PstI site.
Query Match 73.5%; Score 1132.6; DB 1; Length 2568;
Best Local Similarity 98.8%; Pred. No. 3.1e-278;
Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 33 ATAGCTGGTCTGAAATGGCTACTAGATCGCTCTCTGTAATAACAGCCAAATGGTTGTT 92
DB 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCACCAACAGCCAAATAGTTGTT 956
QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGGAAATC 152
DB 957 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGGAAATC 1016
QY 153 GATCTAAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 212
DB 1017 GATCTAAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 1076
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 272
DB 1077 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 1136
QY 273 AAGGCTATTCAAGAACAAATGATCGCTAACGCTCCAGTAACGACGACTTTTGAGGTC 332
DB 1137 AAGGCTATTCAAGAACAAATGATCGCTAACGCTCCAGTAACGACGACTTTTGAGGTC 1196
QY 333 ATTGATTTTGCAGCGCATGCAACATTTACTGATCGAAACGGCAGGCTTACTTTGCTGAC 392
DB 1197 ATTGATTTTGCAGCGCATGCAACATTTACTGATCGAAACGGCAGGCTTACTTTGCTGAC 1256
QY 393 AAGATGTTTGGTAACTTTGCCGACCCCAACCTGTGCCAAGAAATTTTGTCTAAGCGGCAT 452
DB 1257 AAGATGTTTGGTAACTTTGCCGACCCCAACCTGTGCCAAGAAATTTTGTCTAAGCGGCAT 1316
QY 453 GTGCGGTTAGACCATATTAAGAAAAAACAATACAAAAAACAAGGAAATCTGTTGATGTG 512
DB 1317 GTGCGGTTAGACCATATTAAGAAAAAACAATACAAAAAACAAGGAAATCTGTTGATGTG 1376
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGTCTCAAA 572

DB 1377 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGTCTCAAA 1436
QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTTACTA 632
DB 1437 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTTACTA 1496
QY 633 GCTCAAGCACAAGAGCATTTTAAACAAAACACCCAGGCTATAGATTTTATGAACGTGAC 692
DB 1497 GCTCAAGCACAAGAGCATTTTAAACAAAACACCCAGGCTATAGATTTTATGAACGTGAC 1556
QY 693 TCCTCAATCGTCTACTCATGACAAATGACATTTTCGTCAGATTTTACCAATGGATCAAGAG 752
DB 1557 TCCTCAATCGTCTACTCATGACAAATGACATTTTCGTCAGATTTTACCAATGGATCAAGAG 1616
QY 753 TTTTACTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
DB 1617 TTTTACTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1676
QY 813 AATGAGAAATAAACACACCTGACCTCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
DB 1677 AATGAGAAATAAACACACCTGACCTCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1736
QY 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAATACGTT 932
DB 1737 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAATACGTT 1796
QY 933 GATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAAC 992
DB 1797 GATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAAC 1856
QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTCTCTACACAATCTC 1052
DB 1857 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTCTCTACACAATCTC 1916
QY 1053 GATGCTTTTGGTATTATGAGCTATACCTTAACTTGGAAAACTAGAGGATAATCAGATGAC 1112
DB 1917 GATGCTTTTGGTATTATGAGCTATACCTTAACTTGGAAAACTAGAGGATAATCAGATGAC 1976
QY 1113 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGTAC 1172
DB 1977 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGTAC 2036
QY 1173 CATTTAGCTGGTGGT 1187
DB 2037 CATTTAGCTGGTGGT 2051
RESULT 5
A04926
LOCUS A04926 2568 bp DNA linear PAT 15-JUL-1993
DEFINITION S.equisimilis skc gene for streptokinase.
ACCESSION A04926
VERSION A04926.1 GI:412219
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
REFERENCE 1 (bases 1 to 2568)
AUTHORS Ferretti J.J. and Malke, H.
TITLE Streptokinase-coding recombinant vectors
JOURNAL Patent: EP 015137-A 1 14-AUG-1985;
THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
FEATURES
1. .2568
Location/Qualifiers
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
ORIGIN
Query Match 73.5%; Score 1132.6; DB 6; Length 2568;

Best Local Similarity 98.8%; Pred. No. 3.1e-278; Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;			
Qy	33	ATAGCTGGTCCCTGAATGGCTACTAGATCGCTCTCTGTAATAATACGACCAATTTGGTTGTT	92
Db	897	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTTGTT	956
Qy	93	ACGGTGTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC	152
Db	957	AGCGTGTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC	1016
Qy	153	GATCTAATCATCAGCACTGCTCATGAGGAAAGACAGACAGCAAGCTTAAAGTCCAAATCA	212
Db	1017	GATCTAATCATCAGCACTGCTCATGAGGAAAGACAGACAGCAAGCTTAAAGTCCAAATCA	1076
Qy	213	AAACCAATTTGCTACTCATAGTGGCGGATGTCATATAAATCTCAGAAAGCTGACTTACTA	272
Db	1077	AAACCAATTTGCTACTCATAGTGGCGGATGTCATATAAATCTCAGAAAGCTGACTTACTA	1136
Qy	273	AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGTC	332
Db	1137	AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGTC	1196
Qy	333	ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAAACGGCAAGGTCTACTTTGCTGAC	392
Db	1197	ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAAACGGCAAGGTCTACTTTGCTGAC	1256
Qy	393	AAAGATGGTTCGGTACCTTCGCGACCACTGTCACAGAAATTTTTCGTAAGCGGCAT	452
Db	1257	AAAGATGGTTCGGTACCTTCGCGACCACTGTCACAGAAATTTTTCGTAAGCGGCAT	1316
Qy	453	GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCAATCTGTTGATGTG	512
Db	1317	GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCAATCTGTTGATGTG	1376
Qy	513	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTATGACGANTTCAGACAGGTCTCAA	572
Db	1377	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTATGACGANTTCAGACAGGTCTCAA	1436
Qy	573	GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCAAGAAATTA	632
Db	1437	GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCAAGAAATTA	1496
Qy	633	GCTCAAGCAACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGACGTGAC	692
Db	1497	GCTCAAGCAACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGACGTGAC	1556
Qy	693	TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	752
Db	1557	TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	1616
Qy	753	TTTACTTACCGTGTAAATAACCGGAAACAGCTTATAGGATCAATAAAAATCTGCTG	812
Db	1617	TTTACTTACCGTGTAAATAACCGGAAACAGCTTATAGGATCAATAAAAATCTGCTG	1676
Qy	813	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTTAAAAAGGG	872
Db	1677	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTTAAAAAGGG	1736
Qy	873	GAAGAAGCGGTATGATCCCTTTGATCCAGTCACTTGAAACTGTTACCAATCAATACGTT	932
Db	1737	GAAGAAGCGGTATGATCCCTTTGATCCAGTCACTTGAAACTGTTACCAATCAATACGTT	1796
Qy	933	GATGTCGATACCAACGAATTTGCTAAAGGTGAGCAGCTCTTAAACAGTACGCAACGTAAC	992
Db	1797	GATGTCGATACCAACGAATTTGCTAAAGGTGAGCAGCTCTTAAACAGTACGCAACGTAAC	1856
Qy	993	TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTCTTACCAACATCTC	1052
Db	1857	TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTCTTACCAACATCTC	1916
Qy	1053	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGATATACGATGAC	1112

Db	1917	GATGCTTTTGGTATTATGGACTATACCTTAACCTGAAAGAGTAGAGGATAATACGATGAC	1976
Qy	1113	ACCAACGCTATCATACCGCTTTATATGCGGAAGCGACCCGGAAGAGAGAAATGCTAGCTAC	1172
Db	1977	ACCAACGCTATCATACCGCTTTATATGCGGAAGCGACCCGGAAGAGAGAAATGCTAGCTAC	2036
Qy	1173	CATTTAGCTGGTGGT 1187	
Db	2037	CATTTAGCTGGTGGT 2051	
RESULT 6			
SEDEXB	SDSEXB	8931 bp	DNA linear
LOCUS	S.equisimilis dextb, abc, lrp, skc, rel genes and ORF1.		BCT 18-APR-2005
DEFINITION	X72832		
ACCESSION	X72832.1	GI:407876	
VERSION	abc gene; ABC transporter; dextB gene; dextran glucosidase; glucan		
KEYWORDS	1,6-alpha-glucosidase; leucine rich protein; lrp gene; ORF1; rel		
	gene; skc gene; streptokinase; stringent response-like protein.		
	Streptococcus dysgalactiae subsp. equisimilis		
SOURCE	Streptococcus dysgalactiae subsp. equisimilis		
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
	Streptococcus.		
REFERENCE	1 (bases 3621 to 6190)		
AUTHORS	Malke,H., Roe,B. and Ferretti,J.J.		
TITLE	Nucleotide sequence of the streptokinase gene from Streptococcus		
JOURNAL	equisimilis H46A		
PUBMED	Gene 34 (2-3), 357-362 (1985)		
REFERENCE	2 (bases 1 to 4188; 5790 to 8931)		
AUTHORS	Mechohl,U., Steiner,K., Vettermann,S. and Malke,H.		
TITLE	Genetic organization of the streptokinase region of the		
JOURNAL	Streptococcus equisimilis H46A chromosome		
PUBMED	Mol. Gen. Genet. 241 (1-2), 129-140 (1993)		
REFERENCE	3 (bases 1 to 8931)		
AUTHORS	Malke,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-MAY-1993) H. Malke, Institute for Molecular Biology,		
COMMENT	Jena University, Winzerlaer Str 10, 07708 Jena, FRG		
FEATURES	Related sequences: K02986, M19346, X13399 & X13400.		
	Location/Qualifiers		
source	1..8931		
	/organism="Streptococcus dysgalactiae subsp. equisimilis"		
	/mol_type="genomic DNA"		
	/strain="H46A"		
	/isolate="human group C strain"		
	/sub_species="equisimilis"		
	/db_xref="taxon:119602"		
	/chromosome="streptokinase region"		
	/clone_lib="E.coli plasmid library containing subfragments		
	pCWP73, pRH10, pW4"		
	/the submitted seq: pSHD14/16, pSPV19, pSH2, pMF1,		
	complement (89..1761)		
gene	/gene="dextB"		
	complement (89..115)		
terminator	/gene="dextB"		
	complement (136..1749)		
CDS	/gene="dextB"		
	/EC_number="3.2.1.70"		
	/codon_start=1		
	/transl_table=11		
	/product="glucan 1,6-alpha-glucosidase"		
	/protein_id="CAA51348.1"		
	/db_xref="GI:407877"		
	/db_xref="GOA:Q59905"		
	/db_xref="InterPro:IPR006047"		
	/db_xref="InterPro:IPR006589"		
	/db_xref="UniProt/Swiss-Prot:Q59905"		
	/translation="MQQWVHKATVQIPRSPKDTSGNGIGDLKGTISOLDYLOKLG		
	ITAIWLSPTQSPMDNGDYDISYEAIAEIVGNDDMDLLAAANERGIKILMDLVVN		

```

HTSDSHAMFVEARENPSPERDYYIWRDEPNLMSIFSGSAWELDEASGOYYLHLSFK
KQPDLNWENAHYRQIKYIWMNFWIAKGIQGFGRMDVIDLIGKI PDSEITGNGPRLHXYL
KEMNQATFGNHDVMTVGETWGTATPEIARQYSRPENKELSMVFOFHEHVLQHKPNAPKW
DYAEILDVPALATIISKWQTELKLGWNSLFWNNHDLPRVLSIWGNDSTYREKSAKA
LAILLHMRGTYYIYQGEBIGMTNYPFKDLTEVDDIESLNIAKEAMENGVPFAARVMS
IRKGRDNARTPMQSKDTHAGFSEAQETWLPVNPYQEIINADALANQDSIFPTYQQ
LIALRKQDMLYEADYHLLPTADKFPAYQYQFGERTYIVVNVSDQEQVPAKDLAGAE
VVTITVDVKVLETKHLQPDWAFVCVLSV"
RBS      complement(1757..1761)
gene      /gene="dex8"
          complement(1780..3051)
          /gene="abc"
terminator complement(1780..1810)
          /gene="abc"
          /note="hairpin loop"
misc_feature complement(2482..2499)
          /gene="abc"
          /note="Walker motif B"
CDS      complement(2633..2965)
          /gene="abc"
          /codon_start=1
          /transl_table=11
          /product="ABC transporter"
          /protein_id="CAA51349.1"
          /db_xref="GI:600072"
          /db_xref="GOA:O54086"
          /db_xref="InterPro:IPR003439"
          /db_xref="UniProt/TREMBL:Q54086"
          /translation="MVELNLNHYKYPNTHYAVEDFDLDIKKEFIVFVGSGGCK
          STTLRIAGLEDISEGELKIDGEVNDKSPKORDIAMVFQNALYPMVTYDNNAFGL
          KLRKYKRR"
          complement(2644)
          /gene="abc"
          /note="(+) frame shift mutation in H46A"
misc_feature complement(2828..2854)
          /gene="abc"
          /note="Walker motif A"
RBS      complement(2973..2976)
          /gene="abc"
          complement(3026..3031)
          /gene="abc"
          /note="alternative"
          complement(3043..3048)
          /gene="abc"
          complement(3050..3051)
          /gene="abc"
          /note="TG motif"
          complement(3244..4149)
          /gene="lrp"
          complement(3244..3261)
          /gene="lrp"
          /note="hairpin loop"
          complement(3262..4110)
          /gene="lrp"
          /codon_start=1
          /transl_table=11
          /product="leucine rich protein"
          /protein_id="CAA51350.1"
          /db_xref="GI:407878"
          /db_xref="GOA:O54087"
          /db_xref="InterPro:IPR002197"
          /db_xref="InterPro:IPR008931"
          /db_xref="UniProt/Swiss-Prot:Q54087"
          /translation="MELKDYFFEMQVGPPLGDKWVSVKEGDQYVHFKSCLSEKER
          LILLEVGLQGYEVLQPSWPQVLLDHQGNPQLPETSQFIYLNHQVLPADLVLELQ
          QMIAGVILPTSTQTAFLCQATSIKVLRSLGLELPTLESDFGALTMFVGNANVQ
          VAAGTLRSCFECECOLLTAFLKQSGKLLTAENVMLWSILSHQSFPALTTRQHFQFLN
          PQSDMDVVHAWSHGNLVQTAQRLYIHRNSLQYKLDKFAQQSGHLHQLQDLDAFAY
          LFLLLKY"
          complement(3337..3423)
          /gene="lrp"
          /note="leucine zipper motif"
          complement(4117..4120)
RBS

```

```

/gene="lrp"
complement(4141..4146)
/gene="lrp"
complement(4148..4149)
/gene="lrp"
/note="TG motif"
4172..4368
/note="static DNA bending locus"
4392..5837
/gene="skc"
4392..4393
/gene="skc"
/note="TG motif"
4395..4400
/gene="skc"
4401..4406
/note="skc"
/note="alternative"
4428..4431
/gene="skc"
4439..5761
/gene="skc"
/codon_start=1
/transl_table=11
/product="streptokinase"
/protein_id="CAA51351.1"
/db_xref="GI:407879"
/db_xref="GOA:P00779"
/db_xref="InterPro:IPR004093"
/db_xref="InterPro:IPR008124"
/db_xref="UniProt/Swiss-Prot:P00779"
/translation="MKNYLSFGMPALLPALTFTGTVNSVOALAGEWLLDRPSVNSQL
VSVAGIVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSKFPAIDSGAMSHKLEK
ADLLKALQELIANVHNSDDYFEVIDFASDATITDRNGKYFADKDGSVTLPTQVQE
FLLSGHVRVPYKPEQIQONAKSVDEYTVQFTPLNPDDDFRGLKDTLKLTLAIGD
TITSOELLAQAQSLINKNHPGYTIYERDSYVTFVTHDNDIFRILPMDQETVYRVNREQ
AYRKNGSLNEINNTDLJSEKYVYLKKGKPYDPDRSHLKLFTIKYVDVDTNELL
KSQLLTASERNLDFRLYDPRDKAKLLYNNLDAFGIMDYTLTKVKVEDNDDTNRILT
VYMKRPEGENASYHLAYDKDRYTEEREVYSYLRVTGTPTPDNPDK"
4439..4516
/gene="skc"
5796..5837
/gene="skc"
complement(5859..6302)
/note="unnamed protein product; ORF1"
/codon_start=1
/transl_table=11
/protein_id="CAA51352.1"
/db_xref="GI:407880"
/db_xref="GOA:Q54088"
/db_xref="InterPro:IPR003732"
/db_xref="UniProt/Swiss-Prot:Q54088"
/translation="MKLVLRKVASVSDIKGIAGAINQGLLLLVGVGPDAAEDLAY
AVRKIVNMRIFSADQKMNQSIQIDKGSILSVSOFTLYADTKGNRPFTGAAPDMA
SQFYDREQLADFPVVERGVFGADMVSLINDGPTVILDTKCH"
6007..6009
/citation=[1]
/replace="gc"
6162..6164
/citation=[1]
/replace="tg"
complement(6310..6313)
/note="ORF1"
complement(6317..8575)
/gene="rel"
complement(6317..8536)
/gene="rel"
conflict
conflict
RBS
gene
CDS
Query Match 73.5%; Score 1132.6; DB 1; Length 8931;
Best Local Similarity 98.8%; Pred. No. 3.1e-278;
Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 33 ATAGCTGCTCCTGATGCTACTAGATCGTCTTCTGTAATAACAGCAATTGGTTGTT 92

```

Db 4517 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATAGTTGTT 4576
QY 93 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 4577 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 4636
QY 153 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGACGAGCGTTTAACTCCAAATCA 212
Db 4637 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGACGAGCGTTTAACTCCAAATCA 4696
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTTACTA 272
Db 4697 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTTACTA 4756
QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTACAGCAGCTACTTTGAGGTC 332
Db 4757 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTACAGCAGCTACTTTGAGGTC 4816
QY 333 ATTGATTTTTCGAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 392
Db 4817 ATTGATTTTTCGAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 4876
QY 393 AAGATGTTTTCGTTACCTTTCGCGACCCAACTGTCTCAAGAAATTTTTCGTTAAGCGGACAT 452
Db 4877 AAGATGTTTTCGTTACCTTTCGCGACCCAACTGTCTCAAGAAATTTTTCGTTAAGCGGACAT 4936
QY 453 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 4937 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 4996
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 572
Db 4997 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 5056
QY 573 GATACTAAGCTATTGAAACACTAGTCTATCGGTGACACCATCACTCTCAAGAATTACTA 632
Db 5057 GATACTAAGCTATTGAAACACTAGTCTATCGGTGACACCATCACTCTCAAGAATTACTA 5116
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTCAC 692
Db 5117 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTCAC 5176
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTCAGATTTTACCAATGGATCAAGAG 752
Db 5177 TCCTCAATCGTCACTCATGACATGACATTTTCGTCAGATTTTACCAATGGATCAAGAG 5236
QY 753 TTTACTTACCGTTTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 5237 TTTACTTACCGTTTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 5296
QY 813 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTAATCGTCTTAAAAAGGG 872
Db 5297 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTAATCGTCTTAAAAAGGG 5356
QY 873 GAAGAGCGTATGATCCCTTTGATCGCAGTCACTCTGAACTGTTCCACCATCAATAGCTT 932
Db 5357 GAAGAGCGTATGATCCCTTTGATCGCAGTCACTCTGAACTGTTCCACCATCAATAGCTT 5416
QY 933 GATGTCGATACCAACGAAATGCTTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 992
Db 5417 GATGTCGATACCAACGAAATGCTTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 5476
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTTAACTCTTCAACAATCTC 1052
Db 5477 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTTAACTCTTCAACAATCTC 5536
QY 1053 GATGCTTTGGTATTATGACATATACCTTAACTGGAAGAGTAGAGGATATATCAGATGAC 1112
Db 5537 GATGCTTTGGTATTATGACATATACCTTAACTGGAAGAGTAGAGGATATATCAGATGAC 5596
QY 1113 ACCAAGCGTATCATAAACCGTTTATATGAGGCAAGCGACCGAAGGAGGATGCTAGCTAC 1172

Db 5597 ACCAACCGTATCATACACGTTTATATGGGCAAGCGACCCGAGAGAGAAATGCTAGCTAT 5656
QY 1173 CATTTAGCTGGTGGT 1187
Db 5657 CATTTAGCTGGTGGT 5671
RESULT 7
E00522
LOCUS E00522
DEFINITION DNA fragment of plasmid PMF1 into which DNA encoding streptokinase is inserted.
E00522
ACCESSION E00522
VERSION E00522.1 GI:2168801
KEYWORDS JP 1985237995-A/1.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 2568)
Jiyosefu, J.F. and Horusuto, M.
RECOMBINATION VECTOR FOR PRODUCING STREPTOKINASE
TITLE Patent: JP 1985237995-A 1 26-NOV-1985;
JOURNAL PHILIPS PETROLEUM CO
COMMENT OS Streptococcus equisimilis
PN JP 1985237995-A/1
PD 26-NOV-1985
PF 09-OCT-1984 JP 1984212403
PR 10-OCT-1983 DD 83 25523, 02-MAR-1984 US 84 585417 PI
Jiyosefu JIJI FUERTSUTEI, HORUSUTO MARUKE
PC C12N15/00, C12N1/20, C12N9/70, (C12N15/00, C12R1:46), (C12N1/20, PC C12R1:19),
CC (C12N9/70, C12R1:19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Streptococcus equisimilis H46A; CC *source: library=streptococcus equisimilis H46 library; CC *source: clone=lambda L47 skc clone;
FH Key Location/Qualifiers
FT 5'UTR 1..818
FT sig_peptide 819..896
FT /product='streptokinase signal peptide' FT
FT mat_peptide 897..2138
FT /product='streptokinase'
FT CDS 819..2143
FT /product='streptokinase'
FT 3'UTR 2142..2569.
FEATURES
source location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
ORIGIN
Query Match 73.2%; Score 1127.8; DB 6; Length 2568;
Best Local Similarity 98.5%; Pred. No. 5.2e-277;
Matches 1138; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 33 ATAGCTGGTCTGATGGCTACTAGATCGTCTTCTGTAATAATACAGCAATGGTGT 92
Db 897 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGACCAATAGTTGTT 956
QY 93 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 957 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 1016
QY 153 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGACGAGCGTTTAACTCCAAATCA 212
Db 1017 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGACGAGCGTTTAACTCCAAATCA 1076

213 AAACATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTGAGAAAGCTGACTTACTA 272
 1077 AAACATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTGAGAAAGCTGACTTACTA 1136
 273 AAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 332
 1137 AAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 1196
 333 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 392
 1197 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 1256
 393 AAGATGTTGCGGTAACTTTGCGGACCAACCTGTCCCAAGAAATTTTGTAAAGCGGACAT 452
 1257 AAGATGTTGCGGTAACTTTGCGGACCAACCTGTCCCAAGAAATTTTGTAAAGCGGACAT 1316
 453 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
 1317 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1376
 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
 1377 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1436
 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTA 632
 1437 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTA 1496
 633 GCTCAAGCAAAAGCAATTTAAACAAACCAAGGCTATACGATTTATGAACGTGAC 692
 1497 GCTCAAGCAAAAGCAATTTAAACAAACCAAGGCTATACGATTTATGAACGTGAC 1556
 693 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGATCAAGAG 752
 1557 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGATCAAGAG 1616
 753 TTTACTTACCGTGTAAACCTCGGAAACAGCTTTATAGGATCAATAAAATCTGCTCTG 812
 1617 TTTACTTACCGTGTAAACCTCGGAAACAGCTTTATAGGATCAATAAAATCTGCTCTG 1676
 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAGGG 872
 1677 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAGGG 1736
 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTTCAACATCAATACGTT 932
 1737 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTTCAACATCAATACGTT 1796
 933 GATGTCGATACCAACGAATTGCTTAAAGTGAGCAGCTCTTAAACAGTACGCAACGTAAC 992
 1797 GATGTCGATACCAACGAATTGCTTAAAGTGAGCAGCTCTTAAACAGTACGCAACGTAAC 1856
 993 TTAGACTTTAGAGATTTATACGATCTCTGATTAAGGCTAAACTCTCTACAAATCTC 1052
 1857 TTAGACTTTAGAGATTTATACGATCTCTGATTAAGGCTAAACTCTCTACAAATCTC 1916
 1053 GATGCTTTGATTTATGACATATACCTTAACCTGGAAGATAGAGATATCAAGATGAC 1112
 1917 GATGCTTTGATTTATGACATATACCTTAACCTGGAAGATAGAGATATCAAGATGAC 1976
 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAGGAGATGCTAGCTAC 1172
 1977 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAGGAGATGCTAGCTAC 2036
 1173 CATTTAGCTGGTGGT 1187
 2037 CATTTAGCTATGAT 2051

RESULT 8
 CQ797820
 LOCUS
 DEFINITION Sequence 26 from Patent WO2004029256.
 7057 bp DNA linear PAT 20-APR-2004

ACCESSION CQ797820
 VERSION CQ797820.1 GI:46426093
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE other sequences; artificial sequences.
 1
 AUTHORS selman-houssin Soaa,G., aguiar Cabeza,E., gonz lez Quintero,A.D.
 TITLE Vector for the production of transplastomic angiosperm plants
 JOURNAL Patent: WO 2004029256-A 26 08-APR-2004;
 CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
 FEATURES Location/Qualifiers
 source
 1..7057
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /notes="Description of Artificial Sequence: Nucleotide sequence of DNA fragment from the vector pVTPA-Estrep between the rice atpB and tobacco rbcL borders."
 ORIGIN
 Query Match 72.7%; Score 1120.2; DB 6; Length 7057;
 Best Local Similarity 97.6%; Pred. No. 4.5e-275;
 Matches 1137; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 23 ATATACATGATAGTGGTCTCTGAATGGCTACTAGATCGTCTCTCTTAATTAACAGCCA 82
 DB 2013 ATAAACCATGATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCA 2072
 QY 83 ATTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATT 142
 DB 2073 ATTGATTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATT 2132
 QY 143 TTTTGAATTCGATCTAAACATCACGACCTGCTCATGAGGAGAAAGACAGACGCAAGCTTTAAG 202
 DB 2133 TTTTGAATTTGACCTTAACATCACGCTGCTCATGAGGAGAAAGACAGACGCAAGCTTTAAG 2192
 QY 203 TCCAAATCAAAACCAATTTGCTACTGATAGTGGGCGATGTCAATAACTTGAGAAAGC 262
 DB 2193 TCCAAATCAAAACCAATTTGCTACTGATAGTGGGCGATGCCACATATAACTTGAAAGC 2252
 QY 263 TGACTTACTAAAGGCTATTTCAGAAACAAATTGATCGCTTAACGTCACAGTAACGACGACTA 322
 DB 2253 TGACTTACTAAAGGCTATTTCAGAAACAAATTGATCGCTTAACGTCACAGTAACGACGACTA 2312
 QY 323 CTTTGAGGTCATTTGATTTTGCAGAGCATGCAACCATTTACTGATCGAAACGGCAAGGTTCTA 382
 DB 2313 CTTTGAGGTCATTTGATTTTGCAGAGCATGCAACCATTTACTGATCGAAACGGCAAGGTTCTA 2372
 QY 383 CTTTGCTGCAAAAGATGGTTTCGGTAACCTTTGCGGACGCCAACCTGTCCAGAAATTTTGGCT 442
 DB 2373 CTTTGCTGCAAAAGATGGTTTCGGTAACCTTTGCGGACGCCAACCTGTCCAGAAATTTTGGCT 2432
 QY 443 AAGCGGACATGTCGGGCTTAGACCATATAAAGAAACCAATACAAACCAAGCAAGCAATC 502
 DB 2433 AAGCGGACATGTCGGGCTTAGACCATATAAAGAAACCAATACAAACCAAGCAAGCAATC 2492
 QY 503 TGTGTTGTTGGAATATACCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACC 562
 DB 2493 TGTGTTGTTGGAATATACCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACC 2552
 QY 563 AGGTCTCAAAGATATTAAGCTATTGAAACCACTAGCTATCGGTGACCAATCAATCAATCTCA 622
 DB 2553 AGGTCTCAAAGATATTAAGCTATTGAAACCACTAGCTATCGGTGACCAATCAATCAATCTCA 2612
 QY 623 AGAATTTACTAGCTCAAGCACAAGGCAATTTTAAACAAACCAAGCAAGCAATTTACGATTTA 682
 DB 2613 AGAATTTACTAGCTCAAGCACAAGGCAATTTTAAACAAACCAAGCAAGCAATTTACGATTTA 2672
 QY 683 TGAACGTGACTCTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAAT 742
 DB 2673 TGAACGTGACTCTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAAT 2732

KEYWORDS

SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE

AUTHORS Estrada,M.P., Hernandez,L., Perez,A., Rodriguez,P., Serrano,R.,
Rubiera,R., Pedraza,A., Padron,G., Antuch,W., de la Fuente,J. et
al.

TITLE High level expression of streptokinase in Escherichia coli

JOURNAL Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)

PUBMED 1368792

REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsg 115306] from the original journal article.

FEATURES

source
1..1245
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
/note="group: C"
1..1245
/gene="SKC-2"
1..1245
/gene="SKC-2"
/note="streptokinase; Mature protein lacks initial Met"
/codon_start=1
/product="SKC-2"
/protein_id="AAC60418.1"
/db_xref="GI:257197"
/translation="TAGPEWLLDRPSVNNLSQVSVAGTVEGTNQDISLKPFEDITS
RPAHGKTEQGLSPKSPKPPATDSGMPHKLKADLLKATQBLIANVHNSDDYFEVID
FASDATITDRGSKVVFADKQGSVLTPTQVPOFELLSGHVRVPYKEPKIONAKSVDV
EYVQFTPLNPDPPDFRGLKDTLLKTLAIGDTISOEALLAQASILNKTHPGTYIYE
LSSVINDHIDIFRTILPMQBFTHVKNREQAYEINKKSGLEENINNTDLISEKRYV
LKKGKPYDPDFRSLKLFITIKYDVNTNELLSQLLTAERNLDFRDLYPDPRDKAK
LLYNNLDAFGIMDYTLTGKVEDNHDDTNRITVITMGKREGENASVHLAYDKRYTEE
EREVSYLRYTGTPIPDNDK"

ORIGIN

Query Match 72.1%; Score 1111.8; DB 1; Length 1245;

Best Local Similarity 97.7%; Pred. No. 6.5e-273;

Matches 1128; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 33 ATAGCTGTCCTGAATGGCTACTAGATCGTCTTCGTAAATAACAGCAATGGTGT 92
DB 1 ATTGCTGCACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTGT 60
QY 93 AGCGTTGCTGCTGCTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
DB 61 AGCGTTGCTGCTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAAT 120
QY 153 GATCTAAATCAACGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 212
DB 121 GACCTAAATCAACGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 180
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTACTACTA 272
DB 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATATAAATTTGAGAAAGCTACTACTA 240
QY 273 AAGGCTATTCAAGAACAAATTTGATCGTAAACGCTCCACAGTAAACGACACTACTTTGAGTTC 332
DB 241 AAGGCTATTCAAGAACAAATTTGATCGTAAACGCTCCACAGTAAACGACACTACTTTGAGTTC 300
QY 333 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 392
DB 301 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 360
QY 393 AAAGATGTTTGGTAACTTGGCGACCAACCTGTCACAGAAATTTTGTCTAAGCGACAT 452
DB 361 AAAGATGTTTGGTAACTTGGCGACCAACCTGTCACAGAAATTTTGTCTAAGCGACAT 420
QY 453 GTGCGCGTTAGACCATATATAAGAAACCAATATACAAACCAAGCGAAATCTGTGTGATG 512

DB 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
QY 513 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAACAGAGGTCTCAAA 572
DB 481 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAACAGAGGTCTCAAA 540
QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTCATCTCAAGAAATTTACTA 632
DB 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTCATCTCAAGAAATTTACTA 600
QY 633 GCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTTATACGATTTATGAAACGTCAC 692
DB 601 GCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTTATACGATTTATGAAACGTCAC 660
QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 752
DB 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 720
QY 753 TTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
DB 721 TTTTACTTACCAATGTCAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 813 AATGAGAAATAAAACCAACTGACCTGATCTCTGAGAAATATTACGCTTTAAAAAAGGG 872
DB 781 AATGAGAAATAAAACCAACTGACCTGATCTCTGAGAAATATTACGCTTTAAAAAAGGG 840
QY 873 GAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCCACCATCAATAGCTT 932
DB 841 GAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCCACCATCAATAGCTT 900
QY 933 GATGTCGATACCAACCAATTTGCTAAAAAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 992
DB 901 GATGTCACACCAACCAATTTGCTAAAAAGTGACGAGCTCTTAAACAGCTAGCGAACGTAAC 960
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGAGTAAAGCTAAACCTACTCTCAACAATCTC 1052
DB 961 TTAGACTTCAGAGATTTATACGATCTCTGAGTAAAGCTAAACCTACTCTCAACAATCTC 1020
QY 1053 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1112
DB 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1080
QY 1113 ACCAACCGTATCATACCGTTTATATGGCAACGACCTGCTGATGAGGAGGATGCTAGCTAC 1172
DB 1081 ACCAACCGTATCATACCGTTTATATGGCAACGACCTGCTGATGAGGAGGATGCTAGCTAC 1140
QY 1173 CATTTAGCTGGTGGT 1187
DB 1141 CATTTAGCTGGTGGT 1155

RESULT 11

A20015
LOCUS SEQ ID NO: 23; Nucleotide sequence for methionol-streptokinase
DEFINITION fusion protein.
ACCESSION A20015
VERSION A20015.1 GI:1247848
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1257)
AUTHORS PROTEINS AND NUCLEIC ACIDS
TITLE Patent: WO 9109125-A 23 27-JUN-1991;
JOURNAL Location/Qualifiers
FEATURES
source
1..1257
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
4..1251
/codon_start=1
CDS

/transl_table=11
/product="methionyl-streptokinase fusion protein"
/protein_id="CAA01486.1"
/db_xref="GI:1247849"
/translation="MTAGPEWLLDRPSVNNLSQVSVAGTVGNTQDISLKPEIDLT
SPAHGKTEQGLSPKSKPDPATDSGAMPHLEKADLLKALQBLIANVHNDYFEVI
DFASDIITDRNGKVFADKDGVSITLPQBPFLISGHVVRPYKEKPIQONQKSV
VRYTVOFTLNDPDDFRPGLKDTLKTLAGITTSQELLAQSLINHTHPYTIY
ERDSSTVINDNDIFRILPMDQBFYHVNREQAIEINKKSLNEEINNDLISEKY
VLKKGKPYDPDRSHLFTIKYVDVNTNELKSEOLLTASERNLDFRDLXPRDKA
KLLYNLDAFGIMDYTLTGKVEDNDDTNR1IIVYMKRPEGENASHLAYDKDRYTE
EEREVYSILRYTGTPIDPNDK"

ORIGIN

Query Match 72.1%; Score 1111.6; DB 6; Length 1257;
Best Local Similarity 97.5%; Pred. No. 7.3e-273;
Matches 1129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 30 ATGATAGCTGCTCGAATGGCTACTAGATCGTCCTTCTGTAATAACAGCCCAATTGGTT 89
DB 4 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 63

QY 90 GTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTAA 149
DB 64 GTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTAA 123

QY 150 ATCGATCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAA 209
DB 124 ATTAGCTTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAA 183

QY 210 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTA 269
DB 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTA 243

QY 270 CTAAGGCTATTCAAGAACATTTGATCGCTAAACGTCACAGTACAGGACTACTTTGAG 329
DB 244 CTAAGGCTATTCAAGAACATTTGATCGCTAAACGTCACAGTACAGGACTACTTTGAG 303

QY 330 GTCAATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGCGAGCTACTTTGCT 389
DB 304 GTCAATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGCGAGCTACTTTGCT 363

QY 390 GACAAAGATGGTTGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGGA 449
DB 364 GACAAAGATGGTTGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGGA 423

QY 450 CATGTGCGGTAGACCATATAAGAAACCAATAACAAACAGCGAAATCTGTTGAT 509
DB 424 CATGTGCGGTAGACCATATAAGAAACCAATAACAAACAGCGAAATCTGTTGAT 483

QY 510 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTCAGACAGGCTC 569
DB 484 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTCAGACAGGCTC 543

QY 570 AAGATATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 629
DB 544 AAGATATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 603

QY 630 CTAGCTCAAGCAACAAAGCAATTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 689
DB 604 CTAGCTCAAGCAACAAAGCAATTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 663

QY 690 GACTCTCAATCGTCTCATGACAAATGACATTTTCCGTAAGATTTTACCAATGGATCAA 749
DB 664 GACTCTCAATCGTCTCATGACAAATGACATTTTCCGTAAGATTTTACCAATGGATCAA 723

QY 750 GAGTTTACTTACCGTGTAAATCGGACAGCTTATAGGATCAATAAAATCTGGT 809
DB 724 GAGTTTACTTACCGTGTAAATCGGACAGCTTATAGGATCAATAAAATCTGGT 783

QY 810 CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAA 869
DB 784 CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAA 843

870 GGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATAC 929
844 GGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATAC 903

QY 930 GTTGATGTGATACCAACGAATTCCTAAAAGTGAAGAGTCTTTAAACAGCTAGCGAAGCT 989
DB 904 GTTGATGTCAACACCAACGAATTCCTAAAAGTGAAGAGTCTTTAAACAGCTAGCGAAGCT 963

QY 990 AACTTAGACTTTCAGAGATTTATACGATCCCTGCTGATAGGCTAACTCTTCAACAAAT 1049
DB 964 AACTTAGACTTTCAGAGATTTATACGATCCCTGCTGATAGGCTAACTCTTCAACAAAT 1023

QY 1050 CTCGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGGATAATCACCAGT 1109
DB 1024 CTCGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGGATAATCACCAGT 1083

QY 1110 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGC 1169
DB 1084 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGC 1143

QY 1170 TACCAATTTAGCTGGTGGT 1187
DB 1144 TATCAATTTAGCCTATGAT 1161

RESULT 12
I13203
LOCUS I13203 1257 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 25 from patent US 5434073.
ACCESSION I13203
VERSION I13203.1 GI:910551
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1257)
AUTHORS Dawson,K., Hunter,M.G. and Czaplowski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 25 18-JUL-1995;
FEATURES Location/Qualifiers
source 1..1257
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 72.1%; Score 1111.6; DB 6; Length 1257;
Best Local Similarity 97.5%; Pred. No. 7.3e-273;
Matches 1129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 30 ATGATAGCTGCTCGAATGGCTACTAGATCGTCCTTCTGTAATAACAGCCCAATTGGTT 89
DB 4 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 63

QY 90 GTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTAA 149
DB 64 GTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTAA 123

QY 150 ATCGATCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAA 209
DB 124 ATTAGCTTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAA 183

QY 210 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTA 269
DB 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTA 243

QY 270 CTAAGGCTATTCAAGAACATTTGATCGCTAAACGTCACAGTACAGGACTACTTTGAG 329
DB 244 CTAAGGCTATTCAAGAACATTTGATCGCTAAACGTCACAGTACAGGACTACTTTGAG 303

QY 330 GTCAATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGCGAGCTACTTTGCT 389
DB 304 GTCAATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGCGAGCTACTTTGCT 363

QY	390	GACAAAGATGGTTGGTAACTTGGCCGACCCAACTGTCGAAGAAATTTTGTCTAAGCGGA	449
Db	364	GACAAAGATGGTTGGTAACTTGGCCGACCCAACTGTCGAAGAAATTTTGTCTAAGCGGA	423
QY	450	CATGTGCGCGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTGTTGAT	509
Db	424	CATGTGCGCGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTGTTGAT	483
QY	510	GTGGAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGTCTC	569
Db	484	GTGGAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGTCTC	543
QY	570	AAAGATACCTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA	629
Db	544	AAAGATACCTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA	603
QY	630	CTAGCTCAAGCACAAGCAATTTTAAACAAACCCAGCGCTATACGATTTTATGAACGT	689
Db	604	CTAGCTCAAGCACAAGCAATTTTAAACAAACCCAGCGCTATACGATTTTATGAACGT	663
QY	690	GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAA	749
Db	664	GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAA	723
QY	750	GAGTTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGT	809
Db	724	GAGTTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGT	783
QY	810	CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACCGTCTCTTAAAAAA	869
Db	784	CTGAATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACCGTCTCTTAAAAAA	843
QY	870	GGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACTGTGTCAACATCAAATAC	929
Db	844	GGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACTGTGTCAACATCAAATAC	903
QY	930	GTTGATGTCGATACCAACGAATTTGATAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCT	989
Db	904	GTTGATGTCGATACCAACGAATTTGATAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCT	963
QY	990	AACCTAGACTTCAGAGATTTATACGATCCTCGTGTAAAGCTTAAACTCTCTACAACAAT	1049
Db	964	AACCTAGACTTCAGAGATTTATACGATCCTCGTGTAAAGCTTAAACTCTCTACAACAAT	1023
QY	1050	CTCGATGCTTTGGTATTATGAGTATACCTTAACCTGGAAGAGTAGAGGATTAATCAGAT	1109
Db	1024	CTCGATGCTTTGGTATTATGAGTATACCTTAACCTGGAAGAGTAGAGGATTAATCAGAT	1083
QY	1110	GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGC	1169
Db	1084	GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGC	1143
QY	1170	TACCATTTAGCTGGTGGT 1187	
Db	1144	TATCATTTAGCCTATGAT 1161	
RESULT 13			
LOCUS	E01413	1407 bp	DNA linear PAT 29-SEP-1997
DEFINITION	DNA sequence of streptokinase.		
ACCESSION	E01413		
VERSION	E01413.1	GI:2169669	
KEYWORDS	JP 1987296881-A/1.		
SOURCE	Lycopersicon esculentum (Solanum lycopersicum)		
ORGANISM	Lycopersicon esculentum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.		
REFERENCE	1 (bases 1 to 1407)		
AUTHORS	Marii, J.H. and Deebuitsudo, U.S.		
TITLE	YEAST DNA FRAGMENT AND ITS UTILIZATION		
JOURNAL	Patent: JP 1987296881-A 1 24-DEC-1987;		
COMMENT	PHILLIPS PETROLEUM CO PN JP 1987296881-A/1 PD 24-DEC-1987 JP 1987109620 PF 02-MAY-1987 US 86 860960 PR 08-MAY-1986 US 86 860960 PI MARIJ JIEIN HEIJENSON, DEEBUITSUDO UOMATSUKU SUTOROOWAN PC C12N15/00, C12N1/16, C12N9/12, (C12N1/16, C12R1/84); CC strandedness: Double; CC topology: Linear; CC hypothetical: No; CC anti-sense: No; FH Key Location/Qualifiers FH CDS 8..1255 FT /product='streptokinase' FT 5'UTR 1..7 FT 3'UTR 1256..1407. FEATURES source 1..1407 /organism="Lycopersicon esculentum" /mol_type="genomic DNA" /db_xref="taxon:4081" ORIGIN Query Match 72.0%; Score 1110.2; DB 6; Length 1407; Best Local Similarity 97.3%; Pred. No. 1.7e-272; Matches 1130; Conservative 0; Mismatches 33; Indels 0; Gaps 0; QY 25 ATACCATGATGAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTGTGTAATAAATACAGCCCAAT 84 Db 3 ATTCCATGATGTTGGACCTGAGTGGCTGCTAGACGCTCCATCTGTCAACAACAGCCCAAT 62 QY 85 TGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGAATTTAGTCTTTAAATTTT 144 Db 63 TAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGAATTTAGTCTTTAAATTTT 122 QY 145 TTCAATTCGATCTAAATCATCAGACCTGCTCATCGAGGAAGACAGAGCAAGGCTTAAGTC 204 Db 123 TTCAATTCGATCTAAATCATCAGACCTGCTCATCGAGGAAGACAGAGCAAGGCTTAAGTC 182 QY 205 CAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTTGAGAAAGCTG 264 Db 183 CAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTTGAGAAAGCTG 242 QY 265 ACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAACGACGACTACT 324 Db 243 ACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAACGACGACTACT 302 QY 325 TTGAGGTCATTTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACT 384 Db 303 TTGAGGTCATTTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACT 362 QY 385 TTGCTGCAAAAGATGTTGCTGTAACCTTTGCGGACCCAACTGTCGAAAGAAATTTTGTCTAA 444 Db 363 TTGCTGCAAAAGATGTTGCTGTAACCTTTGCGGACCCAACTGTCGAAAGAAATTTTGTCTAA 422 QY 445 GCGGACATGTGCGGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTG 504 Db 423 GCGGACATGTGCGGTTAGACCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTG 482 QY 505 TTGATGTGAATATACCTGATACGTTTACTCCCTTAAACCCCTGATGAGGATTTTACAGCCAG 564 Db 483 TTGATGTGAATATACCTGATACGTTTACTCCCTTAAACCCCTGATGAGGATTTTACAGCCAG 542 QY 565 GTCTCAAGATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAG 624 Db 543 GTCTCAAGATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAG 602 QY 625 AATTACTAGCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATG 684 Db 603 AATTACTAGCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATG 662		

QY 1045 ACAATCTCGATGTTTGGTATTATGGAATACCTTAACCTTAACTGGAAGTAGAGATAATC 1104
DB 1274 ACAATCTCGATGTTTGGTATTATGGAATACCTTAACCTTAACTGGAAGTAGAGATAATC 1333
QY 1105 ACGATGACACCAACCGTATCATACCGTTTATATGGCGAAGCGACCGAGGAGAGATG 1164
DB 1334 ACGATGACACCAACCGTATCATACCGTTTATATGGCGAAGCGACCGAGGAGAGATG 1393
QY 1165 CTAGCTACCAATTTAGCTGGTGGT 1187
DB 1394 CTAGCTATCATTTAGCTATGAT 1416
; RESULT 15
LOCUS I13204 1512 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 27 from patent US 5434073.
ACCESSION I13204
VERSION I13204.1 GI:910552
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1512)
AUTHORS Dawson,K., Hunter,M.G. and Czaplowski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 27 18-JUL-1995;
FEATURES
source location/Qualifiers
1..1512
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 72.0%; Score 1110.2; DB 6; Length 1512;
Best Local Similarity 97.2%; Pred. No. 1.7e-272;
Matches 1130; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
* QY 25 ATACCATGATAGCTGCTGCTGAATGGCTACTAGATCGTCTCTCTGTAATAAATACAGCAAT 84
DB 254 ATAAAGAAATGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCCAAT 313
; QY 85 TGGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 144
DB 314 TAGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 373
QY 145 TTGAATCGATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACGAAGGCTTAAGTC 204
DB 374 TTGAATTTGACCTTAACATCAGACCTGCTCATGGAGGAAAGACAGACGAAGGCTTAAGTC 433
QY 205 CAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTCACATAAATCTTGAGAAAGCTG 264
DB 434 CAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTCACATAAATCTTGAGAAAGCTG 493
QY 265 ACTTACTAAAGGCTATTCAAGAACCAATTTGATCGTAAACGTCACAGTAACGACGACTACT 324
DB 494 ACTTACTAAAGGCTATTCAAGAACCAATTTGATCGTAAACGTCACAGTAACGACGACTACT 553
QY 325 TTGAGGTCATGATTTTGAAGCGATGCAACCATTTACTGATCGAAAACGGAAGGCTTACT 384
DB 554 TTGAGGTCATGATTTTGAAGCGATGCAACCATTTACTGATCGAAAACGGAAGGCTTACT 613
QY 385 TTGCTGACAAAGATGGTTGGTTAACTTCCGACCCCACTGTCCAAAGAAATTTTGGCTAA 444
DB 614 TTGCTGACAAAGATGGTTGGTTAACTTCCGACCCCACTGTCCAAAGAAATTTTGGCTAA 673
QY 445 GCGGACATGTCGCGTTAGACCATATAAAGAAACCAATTAACAAACCAAGCGAAATCTG 504
DB 674 GCGGACATGTCGCGTTAGACCATATAAAGAAACCAATTAACAAACCAAGCGAAATCTG 733
QY 505 TTGATGTGAATATATCTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACCAG 564
DB 734 TTGATGTGAATATATCTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACCAG 793

QY 565 GTCTCAAAAGATACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAG 624
DB 794 GTCTCAAAAGATACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAG 853
QY 625 AATTACTAGCTCTAAGCAACAAGCATTTTAAACAAAACACCCAGGCTATACGATTTATG 684
DB 854 AATTACTAGCTCTAAGCAACAAGCATTTTAAACAAAACCCATCCAGGCTATACGATTTATG 913
QY 685 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGG 744
DB 914 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGG 973
QY 745 ATCAAGAGTTTACTTACCGTGTAAATAACGGAACAAGCTTATAGGATCAATAAAAAAT 804
DB 974 ATCAAGAGTTTACTTACCGTGTCAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAT 1033
QY 805 CTGGTCTGAATGAAGAAAATAAACAACACTGACCTGATCTCTGAGAAATATATACGCTCTTA 864
DB 1034 CTGGTCTGAATGAAGAAAATAAACAACACTGACCTGATCTCTGAGAAATATATACGCTCTTA 1093
QY 865 AAAAAGGGGAAAAGCCGTATGATCCCTTTTGTATCGAGTCACTTGAAACTGTTCAACCATCA 924
DB 1094 AAAAAGGGGAAAAGCCGTATGATCCCTTTTGTATCGAGTCACTTGAAACTGTTCAACCATCA 1153
QY 925 AATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTCTTAAACAGCTAGCG 984
DB 1154 AATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTCTTAAACAGCTAGCG 1213
QY 985 AACGTAACTTAGACTCTTCAGAGATTTATACGATCTCTGATGAAGGCTTAAACTACTCTTACA 1044
DB 1214 AACGTAACTTAGACTCTTCAGAGATTTATACGATCTCTGATGAAGGCTTAAACTACTCTTACA 1273
QY 1045 ACAATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGTAGAGATAATC 1104
DB 1274 ACAATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGTAGAGATAATC 1333
QY 1105 ACGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATG 1164
DB 1334 ACGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATG 1393
QY 1165 CTAGCTACCAATTTAGCTGGTGGT 1187
DB 1394 CTAGCTATCATTTAGCTATGAT 1416

Search completed: February 1, 2006, 04:17:46
Job time : 7836.07 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 931.565 Seconds
(without alignments)
11024.772 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 1541
Sequence: 1 ttgttgaacttaagaagg.....ccttcaccgatgttgtag 1541

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1541	100.0	1541	3	Aaa37644 Chimeric
2	1496.2	97.1	2096	3	Aaa37643 Chimeric
3	1185	76.9	1661	3	Aaa37637 Chimeric
4	1179	76.5	1327	3	Aaa37628 Streptoki
5	1166.2	75.7	1377	3	Aaa37622 Streptoki
6	1134.2	73.6	1782	3	Aaa37642 Chimeric
7	1134	73.6	2385	2	Aax80497 Streptoki
8	1132.8	73.5	2030	2	Aaql1651 FB-FB-SK
9	1132.6	73.5	1242	2	Aax80492 Streptoco
10	1132.6	73.5	1245	3	Aaa37633 S. equisi
11	1132.6	73.5	1254	6	Abao5546 Streptoki
12	1132.6	73.5	8893	6	Abao5547 Maxadlan
13	1131	73.4	1242	2	Aax16632 Streptoco
14	1127.8	73.2	1242	2	Aax16633 Streptoco
15	1120.2	72.7	7057	12	Adm01294 Plasmid p
16	1118.6	72.6	2566	1	Aat77778 Coding se
17	1110.2	72.0	1407	1	Aan70106 DNA encod
18	1110.2	72.0	1512	2	Aaql2158 Streptoki
19	1108.6	71.9	1245	2	Aaq20665 SKC-2 str

20	1108.6	71.9	1335	2	AAQ12156 Streptoki
21	1108.6	71.9	1458	2	AAQ12162 Factor Xa
22	1108.6	71.9	1467	2	AAQ12490 Factor Xa
23	1108.6	71.9	2589	2	AAQ12160 OmpAL str
24	1108.6	71.4	1473	2	AAQ05603 Sequence
25	1094.2	71.0	2568	1	AAN50493 Streptoki
26	1092.6	70.9	1323	2	AAT29961 Vector ps
27	1084.4	70.4	1209	3	AAX99249 DNA encod
28	1084.4	70.4	1245	3	AAX99250 DNA encod
29	1083.8	70.3	1122	3	AAX99251 DNA encod
30	1083.8	70.3	1158	3	AAX99252 DNA encod
31	1083.4	70.3	1242	5	Aaf82144 Mutant st
32	1078.8	70.0	2253	2	AAQ12161 Met-core
33	1077.6	69.9	1119	2	AAQ12159 Truncated
34	993.4	64.5	1320	6	ABN70192 Streptoco
35	993.4	64.5	1323	13	ADR83811 S. pyogen
36	988.6	64.2	1473	2	AAQ05604 Streptoki
37	973	63.1	1068	2	AAX80493 Recombina
38	968.4	62.8	2208	2	AAX83589 Recombina
39	894.6	58.1	1245	10	ADF48644 Streptoco
40	761.2	49.4	1262	2	AAQ10230 Synthetic
41	450	29.2	450	2	AAX80494 Recombina
42	330	21.4	777	3	AAA37632 Human fib
43	330	21.4	1929	13	AD817488 Nucleotid
44	330	21.4	1929	13	ADR97657 Human fib
45	330	21.4	1929	14	ADW44479 Human fib

ALIGNMENTS

RESULT 1
AAA37644
ID AAA37644 standard; DNA; 1541 BP.
XX
AC AAA37644;
XX
15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
Chimeric SK-FBD coding sequence.
XX
Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
EP1024192-A2.
XX
02-AUG-2000.
XX
23-DEC-1999; 99EP-00310541.
XX
24-DEC-1998; 98TN-DE003825.
XX
(COUL) CSIR COUNCIL SCI IND RES.
XX
Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
Yadav M;
XX
WPI; 2000-516032/47.
XX
Hybrid streptokinase-fibrin binding domain polypeptides useful for
thrombolytic therapy comprises a streptokinase fused with fibrin binding
domains of human fibronectin.
XX
Disclosure; Fig 17b; 58pp; English.
XX
This sequence represents a chimeric streptokinase-fibrin binding domain
(SK-FBD) protein coding sequence. The invention relates to a hybrid

plasminogen activator (PA) comprises a polypeptide fusion between streptokinase (SK), which are capable of plasminogen (PG) activation, and fibrin binding regions of human fibrinectin, which are from fibrin binding domains (FBD) 4 and 5 of 1 and 2. The hybrid PA possesses the ability to bind with fibrin independently and also characteristically retains a PG activation ability which becomes evident only after a pronounced duration, or lag, after exposure of the PA to a suitable animal or human PG. The hybrid streptokinase-fibrin binding domain polypeptides are useful in thrombolytic therapy for various kinds of cardiovascular disorders. The hybrids have enhanced fibrin selectivity as well as kinetics of plasminogen activation that are distinct from that of natural streptokinase in being characterised by a temporary delay, or lag of several minutes in the natural rate of the catalytic conversion of plasminogen to plasmin (i.e. delayed-action thrombolytics). The proteins can bind tightly with fibrin in blood clots soon after introduction into the vascular system without significantly activating the circulating blood plasminogen to plasmin, thus aiding in the localisation of the plasminogen activation process to the site of pathological thrombus. This overcomes systemic plasminogen activation encountered during clinical use of streptokinase. (Updated on 15-SEP-2003 to standardise OSfield)

SQ Sequence 154i BP; 497 A; 328 C; 335 G; 381 T; 0 U; 0 Other;

Query Match	100.0%	Score 1541	DB 3	Length 1541
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1541	Conservative	0	Mismatches 0	Indels 0
				Gaps 0

Qy	1	TTTGTTTAACTTTAAGAAGGAGATATACATGATAGCTGGTCTCGTAATGGCTACTAGATC	60
Db	1	TTTGTTTAACTTTAAGAAGGAGATATACCATGATAGCTGGTCTCGTAATGGCTACTAGATC	60
Qy	61	GTCCCTTCGTAAATAACACAGCCAAATTTGGTTGTAGCGGTGCTCGTACTGTGTAGGGGGACCA	120
Db	61	GTCCCTTCGTAAATAACACGCCAAATTTGGTTGTAGCGGTGCTGGTACTGTGTAGGGGGACCA	120
Qy	121	ATCAAGACATTAGTCTTTAAATTTTTTGAATCGATCTAACATCACGACCTGCTCATGGAG	180
Db	121	ATCAAGACATTAGTCTTTAAATTTTTTGAATCGATCTAACATCACGACCTGCTCATGGAG	180
Qy	181	GAAGAAGAGCAAGGCTTTAGTCCAAATCAAAACCATTTGCTACTGATAGTGGGCGCA	240
Db	181	GAAGAAGAGCAAGGCTTTAGTCCAAATCAAAACCATTTGCTACTGATAGTGGGCGCA	240
Qy	241	TGTCACATAAACTTCAGAAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCTA	300
Db	241	TGTCACATAAACTTCAGAAAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTTGATCGCTA	300
Qy	301	ACGTCACACAGTAAACGACGACTACTTTTGAGGTCATTGATTTTCAAGCGGATGCAACCATTA	360
Db	301	ACGTCACACAGTAAACGACGACTACTTTTGAGGTCATTGATTTTCAAGCGGATGCAACCATTA	360
Qy	361	CTGATCGAAACGGCNAAGTCTACTTTGCTGCAAAAGATGGTTCCGGTAACTTGC CGACCC	420
Db	361	CTGATCGAAACGGCNAAGTCTACTTTGCTGCAAAAGATGGTTCCGGTAACTTGC CGACCC	420
Qy	421	AACCTGTCCAAAGATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATAAGAAAAAC	480
Db	421	AACCTGTCCAAAGATTTTGTCTAAGCGGACATGTGCGGTTAGACCATATAAGAAAAAC	480
Qy	481	CAATACAAAAACGACGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA	540
Db	481	CAATACAAAAACGACGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA	540
Qy	541	ACCCCTGATGACGATTTACAGCCAGTCTCAAGATACCTAAGCTATTGAAAAACACTAGCTA	600
Db	541	ACCCCTGATGACGATTTACAGCCAGTCTCAAGATACCTAAGCTATTGAAAAACACTAGCTA	600
Qy	601	TCGGTGACACCATCACATCTCAAGAATTTACTAGCTCAAGCACAAAGCATTTTAAACAAA	660
Db	601	TCGGTGACACCATCACATCTCAAGAATTTACTAGCTCAAGCACAAAGCATTTTAAACAAA	660
Qy	661	ACCAACCCAGGCTATACGATTTATGAACGCTGACTCCTCAATCGCTCACTCATGCAATGACA	720

QY 961 GTGAGCAGCTCTTAACAGCTAGCGAAGCTAACTTAGACTTTCAGAGATTTATACGATCCTC 1020
Db |||||
QY 1111 GTGAGCAGCTCTTAACAGCTAGCGAAGCTAACTTAGACTTTCAGAGATTTATACGATCCTC 1170
Db |||||
QY 1021 GTGATAAGCTAACTACTCTACAACTCTCATGCTTTTGGTATATGGACTATACCT 1080
Db |||||
QY 1171 GTGATAAGCTAACTACTCTACAACTCTCATGCTTTTGGTATATGGACTATACCT 1230
Db |||||
QY 1081 TAACTGGAAGTAGAGGATAATCAGATGACACCAACCGTATCATAAACCGTTTATATGG 1140
Db |||||
QY 1231 TAACTGGAAGTAGAGGATAATCAGATGACACCAACCGTATCATAAACCGTTTATATGG 1290
Db |||||
QY 1141 GCAAGCGACCCGAAGGAGAGAACTGCTAGCTTACCAITTAGCTGGTGGCCAGCGCAAC 1200
Db |||||
QY 1291 GCAAGCGACCCGAAGGAGAGAACTGCTAGCTTACCAITTAGCTGGTGGCCAGCGCAAC 1350
Db |||||
QY 1201 AGATTGTAC 1209
Db |||||
QY 1351 AAATGGTTC 1359
Db |||||

RESULT 4

AAA37628
ID AAA37628 standard; DNA; 1327 BP.

AC AAA37628;

XX
XX
XX

DT 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)

XX
XX
XX

DE Streptokinase-NTR gene.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;

KW plasminogen; human; fibronectin; thrombolytic therapy;

KW cardiovascular disorder; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sabni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

PI Yadav M;

XX WPI; 2000-516032/47.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for

PT thrombolytic therapy comprises a streptokinase fused with fibrin binding

PT domains of human fibronectin.

XX Example 1; Fig 14; 58pp; English.

XX This sequence represents a streptokinase-NTR (SK-NTR) gene (where NTR

CC stands for N-terminally repaired with native sequence). The invention

CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide

CC fusion between streptokinase (SK), which are capable of plasminogen (PG)

CC activation, and fibrin binding regions of human fibronectin, which are

CC from fibrin binding domains (PBD) 4 and 5 or 1 and 2. The hybrid PA

CC possesses the ability to bind with fibrin independently and also

CC characteristically retains a PG activation ability which becomes evident

CC only after a pronounced duration, or lag, after exposure of the PA to a

CC suitable animal or human PG. The hybrid streptokinase-fibrin binding

CC domain polypeptides are useful in thrombolytic therapy for various kinds

CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity

CC as well as kinetics of plasminogen activation that are distinct from that

CC of natural streptokinase in being characterised by a temporary delay, or
CC lag of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX
SQ Sequence 1327 BP; 456 A; 279 C; 250 G; 342 T; 0 U; 0 Other;

Query Match 76.5%; Score 1179; DB 3; Length 1327;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1182; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGCTGGTCCCTGAATGGCTACTAGATC 60
Db |||||
QY 61 GTCCTTCTGTAAATAACAGCCAAATTGGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGA 120
Db |||||
QY 111 GTCCTTCTGTAAATAACAGCCAAATTGGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGA 170
QY 121 ATCAAGACATTAGTCTTAAATTTTTCGAAATCGATCTAAACATCAGCACCTGCTCATGGAG 180
Db |||||
QY 171 ATCAAGACATTAGTCTTAAATTTTTCGAAATCGATCTAAACATCAGCACCTGCTCATGGAG 230
QY 181 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 240
Db |||||
QY 231 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 290
QY 241 TGTCACATAAATCTGAGAAAGCTGACTTACTTAAAGGCTATTCAAGAACATTTGATCGCTA 300
Db |||||
QY 291 TGTCACATAAATCTGAGAAAGCTGACTTACTTAAAGGCTATTCAAGAACATTTGATCGCTA 350
QY 301 ACGTCCACAGTAAACGACGACTTCTTTGAGGTGCTTTGAGGTGCTTTGAGGTGCTTTGAGGTG 360
Db |||||
QY 351 ACGTCCACAGTAAACGACGACTTCTTTGAGGTGCTTTGAGGTGCTTTGAGGTGCTTTGAGGTG 410
QY 361 CTGATCGAAACGGCAAGGCTCTACTTTGCTGACAAAGATGGTTCCGTTAACCTTTGCCGACCC 420
Db |||||
QY 411 CTGATCGAAACGGCAAGGCTCTACTTTGCTGACAAAGATGGTTCCGTTAACCTTTGCCGACCC 470
QY 421 AACCTGTCGAAGAAATTTTTCGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAAC 480
Db |||||
QY 471 AACCTGTCGAAGAAATTTTTCGTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAAC 530
QY 481 CAATACAAAACCAAGCGGAATCTGTTGATGCGGAATATCTGTACAGTATTGTAACCTTACTCCCTAA 540
Db |||||
QY 531 CAATACAAAACCAAGCGGAATCTGTTGATGCGGAATATCTGTACAGTATTGTAACCTTACTCCCTAA 590
QY 541 ACCTGATGACGATTTTCAGACCAAGGCTCTCAAGATATCTAAAGTATTGAAACACCTAGCTA 600
Db |||||
QY 591 ACCTGATGACGATTTTCAGACCAAGGCTCTCAAGATATCTAAAGTATTGAAACACCTAGCTA 650
QY 601 TCGGTGACACCATCAATCTCAAGAAATTTACTAGCTCAAGCAACAAAGCATTTTAAACAAAA 660
Db |||||
QY 651 TCGGTGACACCATCAATCTCAAGAAATTTACTAGCTCAAGCAACAAAGCATTTTAAACAAAA 710
QY 661 ACCACCGAGCTATACGATTTTATGAACGCTGACTCTCTCAATCGTCACATCATGACATGACA 720
Db |||||
QY 711 ACCACCGAGCTATACGATTTTATGAACGCTGACTCTCTCAATCGTCACATCATGACATGACA 770
QY 721 TTTTCCGTGACGATTTTACCAATCGATCAAGAGTTTACTTACCGTGTGTTAAAAATCGGGAAC 780
Db |||||
QY 771 TTTTCCGTGACGATTTTACCAATCGATCAAGAGTTTACTTACCGTGTGTTAAAAATCGGGAAC 830
QY 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAAATGAAGAAATAAAAACAACATGACCTGA 840
Db |||||
QY 831 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAAATGAAGAAATAAAAACAACATGACCTGA 890
QY 841 TCTCTGAGAAATATATAGTCTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTTGTATCGCA 900

Db 891 TCTCTGAGAAATATTACGTCCTTTAAAAAAGGGGAAAAAGCGTATGATCCCTTTGATCGCA 950
 Qy 901 GTCACCTTGAACCTGTTTACCACCAATAGCTTGCATGCTGCATACCAACGAAATGCTAATA 960
 Db 951 GTCACCTTGAACCTGTTTACCACCAATAGCTTGCATGCTGCATACCAACGAAATGCTAATA 1010
 Qy 961 GTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTAGACTTTCAGAGATTTATAGATCCCTC 1020
 Db 1011 GTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTAGACTTTCAGAGATTTATAGATCCCTC 1070
 Qy 1021 GTGATAAGGCTAAACCTTACTCTACCAACATCTCGATGCTTTTGGTATTTATGGAATACCT 1080
 Db 1071 GTGATAAGGCTAAACCTTACTCTACCAACATCTCGATGCTTTTGGTATTTATGGAATACCT 1130
 Qy 1081 TTAACCTGGAACCTAGAGGATATCAGATGACACCAACCGTATCATACCGTTTATATGG 1140
 Db 1131 TTAACCTGGAACCTAGAGGATATCAGATGACACCAACCGTATCATACCGTTTATATGG 1190
 Qy 1141 GCAAGCGACCCGAAGGAGAGAAATGCTAGCTACCATTAGCTGCTGGT 1187
 Db 1191 GCAAGCGACCCGAAGGAGAGAAATGCTAGCTATCATTTAGCCTATGAT 1237

RESULT 5

AAA37622
 ID AAA37622 standard; DNA; 1377 BP.
 XX
 AC AAA37622;
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE Streptokinase-NTRN gene.
 XX
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibronectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 PN EP1024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RES.
 XX
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX
 DR WPI; 2000-516032/47.
 XX
 PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibronectin.
 XX
 PS Example 1; Fig 11; 58pp; English.
 XX
 CC This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN
 CC stands for N-terminally repaired with native sequence). The invention
 CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
 CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
 CC activation, and fibrin binding regions of human fibronectin, which are
 CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
 CC possesses the ability to bind with fibrin independently and also
 CC characteristically retains a PG activation ability which becomes evident
 CC only after a pronounced duration, or lag, after exposure of the PA to a
 CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
 CC domain polypeptides are useful in thrombolytic therapy for various kinds

CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
 CC as well as kinetics of plasminogen activation that are distinct from that
 CC of natural streptokinase in being characterised by a temporary delay, or
 CC lag of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX

SQ Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;

Query Match 75.7%; Score 1166.2; DB 3; Length 1377;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1174; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Qy 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 60
 Db 101 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 160
 Qy 61 GTCCTTCTGTAATAACAGCCAAATGGTTTGTAGCGTTGCTGCTACTGTTGAGGGGACGA 120
 Db 161 GTCCTTCTGTAATAACAGCCAAATGGTTTGTAGCGTTGCTGCTACTGTTGAGGGGACGA 220
 Qy 121 ATCAAGACATTAGTCTTAAATTTTGGAAATCGATTAACATCAAGACCTGCTCATGAG 180
 Db 221 ATCAAGACATTAGTCTTAAATTTTGGAAATCGATTAACATCAAGACCTGCTCATGAG 280
 Qy 181 GAAAGACAGAGCAAGCTTAAAGTCCAAATCAAAATCCATTTGCTACTGATAGTGGCCGA 240
 Db 281 GAAAGACAGAGCAAGCTTAAAGTCCAAATCAAAATCCATTTGCTACTGATAGTGGCCGA 340
 Qy 241 TGTCACATAAACTTGAGAAAGCTGACTTACTAAAGGCTATTTCAAGACAAATGATCGCTA 300
 Db 341 TGTCACATAAACTTGAGAAAGCTGACTTACTAAAGGCTATTTCAAGACAAATGATCGCTA 400
 Qy 301 ACGTCCACAGTAACGACGACTACTTTGAGTCAATGATTTTGAAGGATGCAACCAATTA 360
 Db 401 ACGTCCACAGTAACGACGACTACTTTGAGTCAATGATTTTGAAGGATGCAACCAATTA 460
 Qy 361 CTGATCGAAACGCGAAGGCTCTACTTTGCTGACAAAGATGGTTGCGTAACCTTTGCCGACC 420
 Db 461 CTGATCGAAACGCGAAGGCTCTACTTTGCTGACAAAGATGGTTGCGTAACCTTTGCCGACC 520
 Qy 421 AACCTGTCCAAAGAAATTTTGTAAAGCGGACATGTGCGGCTTAGACCATATAAAGAAAAAC 480
 Db 521 AACCTGTCCAAAGAAATTTTGTAAAGCGGACATGTGCGGCTTAGACCATATAAAGAAAAAC 580
 Qy 481 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 540
 Db 581 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAA 640
 Qy 541 ACCCTGATGACGATTTTCAGACAGGCTCTCAAGATCTAAGCTATTTGAAACACTAGCTA 600
 Db 641 ACCCTGATGACGATTTTCAGACAGGCTCTCAAGATCTAAGCTATTTGAAACACTAGCTA 700
 Qy 601 TCGGTGACACCATCAATCTCAAGAAATTTACTAGCTCAAGCAACAAAGCATTTTAAACAAA 660
 Db 701 TCGGTGACACCATCAATCTCAAGAAATTTACTAGCTCAAGCAACAAAGCATTTTAAACAAA 760
 Qy 661 ACCACCGAGCTATACGATTTATGAACGCTGACTCTCAATCGTCACTCATGACAAATGACA 720
 Db 761 ACCACCGAGCTATACGATTTATGAACGCTGACTCTCAATCGTCACTCATGACAAATGACA 820
 Qy 721 TTTTCCGTAGCATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 780
 Db 821 TTTTCCGTAGCATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 880
 Qy 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 840
 Db 881 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 940

QY 841 TCTCTGAGAAATATTACGTCCTTTAAATAAGGGGAAAGCGGTATGATCCCTTTGATGCGCA 900
Db TCTCTGAGAAATATTACGTCCTTTAAATAAGGGGAAAGCGGTATGATCCCTTTGATGCGCA 1000
QY 901 GTCACTTGAACCTGTTACCATCAATACGTTGATGTCGATACCAACGAATTCCTAAAAA 960
Db 1001 GTCACTTGAACCTGTTACCATCAATACGTTGATGTCGATACCAACGAATTCCTAAAAA 1060
QY 961 GTGAGCAGCTCTTAAACAGCTAGCGAACTTAACTTACAGACTTTCAGAGATTTATACGATCCTC 1020
Db 1061 GTGAGCAGCTCTTAAACAGCTAGCGAACTTAACTTACAGACTTTCAGAGATTTATACGATCCTC 1120
QY 1021 GTGATAAGGCTAACTCTTACACAATCTCGATGCTTTTGGTATTTATGAGACTATACCT 1080
Db 1121 GTGATAAGGCTAACTCTTACACAATCTCGATGCTTTTGGTATTTATGAGACTATACCT 1180
QY 1081 TAACTGGAAGTAGAGGATAATCAGATGACACCAACCGTATCATAAACCGTTTATATGG 1140
Db 1181 TAACTGGAAGTAGAGGATAATCAGATGACACCAACCGTATCATAAACCGTTTATATGG 1240
QY 1141 GCAAGCGACCCGAGGAGAGAACTAGCTAGCTTACCATTTAGCTGGT 1187
Db 1241 GCAAGCGACCCGAGGAGAGAACTAGCTAGCTTACCATTTAGCTATGAT 1287

RESULT 6

AAA37642

ID AAA37642 standard; DNA; 1782 BP.

XX

AC AAA37642;

XX 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)

XX

Chimeric SK-FBD coding sequence.

DE

KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;

KW plasminogen; human; fibrinectin; thrombolytic therapy;

KW cardiovascular disorder; ss.

XX

Streptococcus dysgalactiae subsp. equisimilis.

OS Homo sapiens.

OS Chimeric.

FN EP1024192-A2.

XX

PD 02-AUG-2000.

XX

PF 23-DEC-1999; 99EP-00310541.

XX

PR 24-DEC-1998; 98IN-DE003825.

XX

(COUL) CSIR COUNCIL SCI IND RES.

PA

Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

PI Yadav M;

XX

WPI; 2000-516032/47.

DR

Hybrid streptokinase-fibrin binding domain polypeptides useful for thrombolytic therapy comprises a streptokinase fused with fibrin binding domains of human fibrinectin.

PT

Example 5; Fig 21b; 58pp; English.

PS

This sequence represents a chimeric streptokinase-fibrin binding domain (SK-FBD) protein coding sequence. The invention relates to a hybrid plasminogen activator (PA) comprises a polypeptide fusion between streptokinase (SK), which are capable of plasminogen (PG) activation, and fibrin binding regions of human fibrinectin, which are from fibrin binding domains (PBD) 4 and 5 or 1 and 2. The hybrid PA possesses the ability to bind with fibrin independently and also characteristically

CC retains a PG activation ability which becomes evident only after a pronounced duration, or lag, after exposure of the PA to a suitable animal or human PG. The hybrid streptokinase-fibrin binding domain polypeptides are useful in thrombolytic therapy for various kinds of cardiovascular disorders. The hybrids have enhanced fibrin selectivity as well as kinetics of plasminogen activation that are distinct from that of natural streptokinase in being characterised by a temporary delay, or lag of several minutes in the natural rate of the catalytic conversion of plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins can bind tightly with fibrin in blood clots soon after introduction into the vascular system without significantly activating the circulating blood plasminogen to plasmin, thus aiding in the localisation of the plasminogen activation process to the site of pathological thrombus. This overcomes systemic plasminogen activation encountered during clinical use of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX

SQ Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Query Match 73.6%; Score 1134.2; DB 3; Length 1782;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 33 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTCTGTGTAATAACAGCAATTCGGTGT 92

Db 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTCGGTGT 597

QY 93 AGCGTCTGCTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152

Db 598 AGCGTCTGCTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 657

QY 153 GATCTAACATCAGCACCTCTGATGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 212

Db 658 GATCTAACATCAGCACCTCTGATGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 717

QY 213 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAGCTGACTTACTA 272

Db 718 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAGCTGACTTACTA 777

QY 273 AAGCTATTCAAGAACAAATTGATCGTCAAGTCCACAGTAACGACGACTTTGAGGTC 332

Db 778 AAGCTATTCAAGAACAAATTGATCGTCAAGTCCACAGTAACGACGACTTTGAGGTC 837

QY 333 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCGAAGTCTTCTTTGCTGAC 392

Db 838 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCGAAGTCTTCTTTGCTGAC 897

QY 393 AAAGATGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 452

Db 898 AAAGATGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 957

QY 453 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 512

Db 958 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 1017

QY 513 GAATATCTCTACAGTTTACTCCCTTAAACCCCTGATCGAATTTTCAAGCAGGCTCTCAA 572

Db 1018 GAATATCTCTACAGTTTACTCCCTTAAACCCCTGATCGAATTTTCAAGCAGGCTCTCAA 1077

QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 632

Db 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 1137

QY 633 GCTCAAGCAACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 692

Db 1138 GCTCAAGCAACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC 1197

QY 693 TCCTCAATCTGCTACTCATGACAAATGACATTTTCGTAGGATTTTACCAATGGATCAAGAG 752

Db 1198 TCCTCAATCTGCTACTCATGACAAATGACATTTTCGTAGGATTTTACCAATGGATCAAGAG 1257

QY 753 TTTTACTTACCGTGTATAAATTCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Db 1258 TTTTACTTACCGTGTATAAATTCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317

QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
| | | | |
Db 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
| | | | |
QY 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTACCAATCAAAATACGTT 932
| | | | |
Db 1378 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTACCAATCAAAATACGTT 1437
| | | | |
QY 933 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAAC 992
| | | | |
Db 1438 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAAC 1497
| | | | |
QY 993 TTAGACTTCAGAGATTATACGATCTCTGATGATGAGGCTAAAGCTACTCTACCAATCTC 1052
| | | | |
Db 1498 TTAGACTTCAGAGATTATACGATCTCTGATGATGAGGCTAAAGCTACTCTACCAATCTC 1557
| | | | |
QY 1053 GATGCTTTTGGTATTTATGACATATACCTTTAACTGGAAGAGTAGAGGATAATCACGATGAC 1112
| | | | |
Db 1558 GATGCTTTTGGTATTTATGACATATACCTTTAACTGGAAGAGTAGAGGATAATCACGATGAC 1617
| | | | |
QY 1113 ACCAACCGTATACATACCGTTTATATGGCGAAGCGACCGAAGGAGAGATGCTAGTAC 1172
| | | | |
Db 1618 ACCAACCGTATACATACCGTTTATATGGCGAAGCGACCGAAGGAGAGATGCTAGTAT 1677
| | | | |
QY 1173 CATTTAGCTGGTGGT 1187
| | | | |
Db 1678 CATTTAGCTATGAT 1692
| | | | |

RESULT 7

AAx80497
ID AAx80497 standard; cDNA; 2385 BP.
XX
AC AAx80497;
XX
DT 26-AUG-1999 (first entry)
XX
DE Streptokinase and maltose binding protein fusion protein encoding cDNA.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nSK; rSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
XX
XX WO9931247-A1.
XX
XX 24-JUN-1999.
XX
XX 15-DEC-1998; 98WO-US026694.
XX
XX 15-DEC-1997; 97US-0069497P.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Reed GL;
XX
XX WPI; 1999-395183/33.
XX
XX P-PSDB; AAY24797.
XX
XX N-terminally deleted streptokinase.
XX
XX Example; Page 45-48; 73pp; English.
XX
XX
XX The present invention describes an isolated bacterial protein that
XX induces fibrin-dependent plasminogen activation in a pharmaceutical
XX composition for dissolving blood clots. Also described are: (1) a
XX composition comprising an isolated modified streptokinase, the
XX modification being removal of amino acid residues in the amino terminus;
XX (2) a method for dissolving a blood clot in a subject, comprising

CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a
CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC blood clots in patients with a thrombotic condition, e.g. myocardial
CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
CC graft thrombosis and arterial thrombosis. The modified streptokinase can
CC also be used in non-human mammals. Streptokinase activation of
CC plasminogen is at least 10-fold, preferably 100-fold greater in the
CC presence of fibrin than in the absence of fibrin. The modified
CC streptokinase has at least one amino acid substitution that inactivates a
CC substrate site for proteolytic cleavage. This reduces the rate of
CC degradation of the streptokinase at least two-fold. The present sequence
CC encodes a streptokinase and maltose binding protein fusion protein from
CC an example of the present invention
XX
SQ Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;
Query Match 73.6%; Score 1134; DB 2; Length 2385;
Best Local Similarity 98.7%; Pred.No. 0;
Matches 1143; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 30 ATGATAGCTGGTCTCTGAATGGCTACTAGATCGTCCTTCTCTAAATACAGCCAAATGGTT 89
| | | | |
Db 1141 AGGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCAAATAGTT 1200
| | | | |
QY 90 GTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAA 149
| | | | |
Db 1201 GTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAA 1260
| | | | |
QY 150 ATCGATCTAACATCAGACCTGCTCATGGAGGAAGACAGACGAAGCTTTAAGTCCAAAA 209
| | | | |
Db 1261 ATCGATCTAACATCAGACCTGCTCATGGAGGAAGACAGACGAAGCTTTAAGTCCAAAA 1320
| | | | |
QY 210 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAGCTGACTTA 269
| | | | |
Db 1321 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAGCTGACTTA 1380
| | | | |
QY 270 CTAAAGGCTATTCAAGAACAAATTTGATCGTCAAGTCAAGTAAACGACGACTCTTTGAG 329
| | | | |
Db 1381 CTAAAGGCTATTCAAGAACAAATTTGATCGTCAAGTCAAGTAAACGACGACTCTTTGAG 1440
| | | | |
QY 330 GTCAATGATTTTGCAGAGCGATGCAACCATTTACTGATCGAAACGCGAGGCTCTACTTTGCT 389
| | | | |
Db 1441 GTCAATGATTTTGCAGAGCGATGCAACCATTTACTGATCGAAACGCGAGGCTCTACTTTGCT 1500
| | | | |
QY 390 GACAAAGATGGTTTCGGTAACTTTGCCGACCCAACTGTGCCAAGAAATTTTGTCTAAGCGGA 449
| | | | |
Db 1501 GACAAAGATGGTTTCGGTAACTTTGCCGACCCAACTGTGCCAAGAAATTTTGTCTAAGCGGA 1560
| | | | |
QY 450 CATGTCGGCTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 509
| | | | |
Db 1561 CATGTCGGCTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 1620
| | | | |
QY 510 GTGGATATATCTGTACAGTTTACTCCCTTAAACCTGTGACGATCCGATTTTCAGACGAGTCTC 569
| | | | |
Db 1621 GTGGATATATCTGTACAGTTTACTCCCTTAAACCTGTGACGATTTTCAGACGAGTCTC 1680
| | | | |
QY 570 AAAGATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 629
| | | | |
Db 1681 AAAGATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1740
| | | | |
QY 630 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACACCCAGGCTATACGATTTATGAACGT 689
| | | | |
Db 1741 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACACCCAGGCTATACGATTTATGAACGT 1800
| | | | |
QY 690 GACTCTCTCAATCGTCACTCATGCAATGACATTTTCGTTACGATTTTACCAATGGATCAA 749
| | | | |
Db 1801 GACTCTCTCAATCGTCACTCATGCAATGACATTTTCGTTACGATTTTACCAATGGATCAA 1860
| | | | |
QY 750 GAGTTTACTTTACCGTGTAAAAAATCGGAACAGCTTTATAGGATCAATAAAAAATCTGGT 809
| | | | |

Db 1861 GAGTTTACTTACCGTGTAAATAACCGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 1920
QY 810 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAA 869
Db 1921 CTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAA 1980
QY 870 GGGGAAAGCCGTATGATCCCTTTGATCGAGTCACCTTGAAACTGTTTCAACATCAATAC 929
Db 1981 GGGGAAAGCCGTATGATCCCTTTGATCGAGTCACCTTGAAACTGTTTCAACATCAATAC 2040
QY 930 GTTGATGTGATACCAACGAATTGCTAAAGTGAAGCGAGCTCTTAACAGCTAGCGACGT 989
Db 2041 GTTGATGTGATACCAACGAATTGCTAAAGTGAAGCGAGCTCTTAACAGCTAGCGACGT 2100
QY 990 AACTTAGACTTCAGAGATTTATACGATCCTCTGCTGATAAGGCTAAACTTACTCTACAACAT 1049
Db 2101 AACTTAGACTTCAGAGATTTATACGATCCTCTGCTGATAAGGCTAAACTTACTCTACAACAT 2160
QY 1050 CTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAGTGAAGGATAATCACGAT 1109
Db 2161 CTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAGTGAAGGATAATCACGAT 2220
QY 1110 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAAGAGAGATGCTAGC 1169
Db 2221 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAAGAGAGATGCTAGC 2280
QY 1170 TACCATTTAGCTGGTGT 1187
Db 2281 TATCATTTAGCCTATGAT 2298

RESULT 8

ID AAQ11651 standard; DNA; 2030 BP.
XX AAQ11651;
AC AAQ11651;
DT 08-JUL-1991 (first entry)
XX
DE FB-PB-SK fusion construct.
XX
KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW streptokinase; fusion protein; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 10..184
FT FT /*tag= a
FT FT /label= FB monomer
FT FT 185..358
FT FT /*tag= b
FT FT /label= FB monomer
FT FT 359..1601
FT FT /*tag= c
FT FT /label= streptokinase
XX
PN US5011686-A.
XX
PD 30-APR-1991.
XX
PF 15-NOV-1989; 89US-00437769.
XX
PR 21-SEP-1987; 87US-00099242.
XX
PA (CREA-) CREATIVE BIOMOLEC.
XX
PI Pang RHL;
XX
XX WPI; 1991-140198/19.
DR P-PSDB; AAR11829.
XX
PT Imparting injectable fibrinolytic agent - with affinity for intravascular

thrombus, by linking agent to fibrin binding domain.
XX Disclosure; Fig 5; 18pp; English.
XX
CC The DNA encodes an FB-PB dimer linked to the streptokinase coding
sequence. The FB fragment has selective affinity for fibrin, low affinity
CC for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
CC capability. See also AAQ11649 and AAQ11650
XX
SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;
Query Match 73.5%; Score 1132.8; DB 2; Length 2030;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 28 CCATGATAGCTGGTCCCTGAATGGCTACTAGATCGTCTCTCTGTAATAACAGCAATGG 87
Db 353 CGAAGATTGCTGGACCTGGAGTGGCTGCTAGACCGTCCATCTGTCAACAACACCAATTAG 412
QY 88 TTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGG 147
Db 413 TTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGG 472
QY 148 AAATCGATCTAAACATCAACGACCTGCTCATGGAGGAAAGACAGACAGAGGCTTAAGTCAA 207
Db 473 AAATCGATCTAAACATCAACGACCTGCTCATGGAGGAAAGACAGACAGAGGCTTAAGTCAA 532
QY 208 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATCTTGAGAAAGCTGACT 267
Db 533 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATCTTGAGAAAGCTGACT 592
QY 268 TACTAAAGGCTATTCAAGAACAAATTGATCGTCAAGTCCACAGTAAACGACGACTACTTTG 327
Db 593 TACTAAAGGCTATTCAAGAACAAATTGATCGTCAAGTCCACAGTAAACGACGACTACTTTG 652
QY 328 AGGTCATTGATTTTGCAGAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTG 387
Db 653 AGGTCATTGATTTTGCAGAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTG 712
QY 388 CTGACAAAGATGTTTGGTTAACTTGCAGCCCAACCTGTCCAAGAAATTTTGTAAAGCG 447
Db 713 CTGACAAAGATGTTTGGTTAACTTGCAGCCCAACCTGTCCAAGAAATTTTGTAAAGCG 772
QY 448 GACATGTCGCGTTAGACCATATAAGAAAAAACAATAAACAACCAAGCAATCTGTTG 507
Db 773 GACATGTCGCGTTAGACCATATAAGAAAAAACAATAAACAACCAAGCAATCTGTTG 832
QY 508 ATGTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGATGATTTTCAGACCGAGTC 567
Db 833 ATGTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGATGATTTTCAGACCGAGTC 892
QY 568 TCAAAGATACTAAGCTATTGAAACACACTAGCTATCGGTGACCACTCACATCTCAAGAAT 627
Db 893 TCAAAGATACTAAGCTATTGAAACACACTAGCTATCGGTGACCACTCACATCTCAAGAAT 952
QY 628 TACTAGCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACATTTATGAAC 687
Db 953 TACTAGCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACATTTATGAAC 1012
QY 688 GTGACTCCTCAATCGTCACTCATGCAATGACATTTTCCGTAGATTTTACCAGTGCATC 747
Db 1013 GTGACTCCTCAATCGTCACTCATGCAATGACATTTTCCGTAGATTTTACCAGTGCATC 1072
QY 748 AAGAGTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCTG 807
Db 1073 AAGAGTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCTG 1132
QY 808 GTCTGAATGAAGAAATAAACAACACTGACTGATCTCTGAGAAATATTACGTCCTTAAAA 867
Db 1133 GTCTGAATGAAGAAATAAACAACACTGACTGATCTCTGAGAAATATTACGTCCTTAAAA 1192
QY 868 AAGGGAAAAAGCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCCACCATCAAT 927

Db 1193 AAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTCCACCATCAAAAT 1252
 QY 928 AGCTTGATGTCGATACCAACGAATTCGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAC 987
 Db 1253 AGCTTGATGTCGATACCAACGAATTCGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAC 1312
 QY 988 GTAACCTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAACA 1047
 Db 1313 GTAACCTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAACA 1372
 QY 1048 ATCTCGATGCTTTGGTATATATGAGACTATACCTTAACCTGGAAGTAGAGATAATCAAG 1107
 Db 1373 ATCTCGATGCTTTGGTATATATGAGACTATACCTTAACCTGGAAGTAGAGATAATCAAG 1432
 QY 1108 ATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGGAGAGAGAAATGCTA 1167
 Db 1433 ATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGGAGAGAGAAATGCTA 1492
 QY 1168 GCTACCATTTAGCTGGT 1187
 Db 1493 GCTATCATTTAGCTATGAT 1512

RESULT 9

AA80492
 ID AAX80492 standard; cDNA; 1242 BP.
 XX AC AAX80492;
 XX 17-OCT-2003 (revised)
 DT 26-AUG-1999 (first entry)
 XX Streptococcus equisimilis native streptokinase encoding cDNA.
 XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW nSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX WO9931247-A1.
 FN 24-JUN-1999.
 PD 15-DEC-1998; 98WO-US026694.
 PF 15-DEC-1997; 97US-0069497P.
 PR (HARD) HARVARD COLLEGE.
 XX Reed GL;
 PI WPI; 1999-395183/33.
 DR P-PSDB; AAY24794.
 XX N-terminally deleted streptokinase.
 PT Claim 44; Page 58-60; 73pp; English.

The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an expression vector comprising (1); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,

CC graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence encodes native streptokinase (nSK). (Updated on 17-OCT-2003 to standardise OS field)
 XX Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;
 SQ Query Match 73.5%; Score 1132.6; DB 2; Length 1242;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCTCGAATGGCTTACGATCGCTCTTGTAAATAACAGCCAAATGGTGT 92
 Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTGT 60
 QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 152
 Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 120
 QY 153 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTTAAGTCCAAATCA 212
 Db 121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTTAAGTCCAAATCA 180
 QY 213 AAACCAATTTGCTACTGATAGTGGCGCATGTCTACATAAACTTGAGAAGCTGACTACTA 272
 Db 181 AAACCAATTTGCTACTGATAGTGGCGCATGTCTACATAAACTTGAGAAGCTGACTACTA 240
 QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAAACGACGACTTTGAGTTC 332
 Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAAACGACGACTTTGAGTTC 300
 QY 333 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGAAACCGCAAGGCTCTACTTTGCTGAC 392
 Db 301 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGAAACCGCAAGGCTCTACTTTGCTGAC 360
 QY 393 AAAGATGGTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
 Db 361 AAAGATGGTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 420
 QY 453 GTGCGGTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTGTGATGT 512
 Db 421 GTGCGGTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTGTGATGT 480
 QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGCTCTCAA 572
 Db 481 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGCTCTCAA 540
 QY 573 GATACTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 632
 Db 541 GATACTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 600
 QY 633 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTTATACGATTTATGAACGTCAC 692
 Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTTATACGATTTATGAACGTCAC 660
 QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
 Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
 QY 753 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 721 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
 QY 813 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 872
 Db 781 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTAACGTCCTTAAAAAGGG 840
 QY 873 GAAAGCCGTATGATCCCTTTTGTATGCGACGTCATCTTGAACCTGTTCCACCATCAATAGT 932

841 GAAAGCCGATGATCCCTTTGATCGAGTCACCTTGAACCTGTTCAACATCAATAGCTT 900
933 GATGTCGATACCAACGAATGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAC 992
901 GATGTCGATACCAACGAATGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAC 960
993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1052
961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
1053 GATGCTTTGGTATTATAGCACTATACCTTAACTCGGAAAAGTAGAGGATAATCAGATGAC 1112
1021 GATGCTTTGGTATTATAGCACTATACCTTAACTCGGAAAAGTAGAGGATAATCAGATGAC 1080
1113 ACCAACCGTATCATACCGTTTATATGGCAAGCGCCGAGGAGAGAGATGCTAGCTAC 1172
1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGCCGAGGAGAGAGATGCTAGCTAT 1140
1173 CATTTAGCTGGTGGT 1187
1141 CATTTAGCTATGAT 1155

RESULT 10

AAA37633
ID AAA37633 standard; DNA; 1245 BP.

AC AAA37633;

DT 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)

XX S. equisimilis streptokinase coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; fibrinectin; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers

FT 1. .1245

FT CDS /*tag= a

FT /product= "streptokinase"

FT EP1024192-A2.

XX PD 02-AUG-2000.

XX PF 23-DEC-1999; 99EP-00310541.

XX PR 24-DEC-1998; 98IN-DE003825.

XX PA (COUL) CSIR COUNCIL SCI IND RES.

XX PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

XX PI Yadav M;

XX DR WPI; 2000-516032/47.

XX DR P-PSDB; AAY90282.

XX XX Hybrid streptokinase-fibrin binding domain polypeptides useful for

XX XX thrombolytic therapy comprises a streptokinase fused with fibrin binding

XX XX domains of human fibrinectin.

XX XX Example 3; Fig 3; 58pp; English.

XX CC This sequence represents the human Streptococcus equisimilis

XX CC streptokinase coding sequence. The invention relates to a hybrid

XX CC plasminogen activator (PA) comprises a polypeptide fusion between

XX CC streptokinase (SK), which are capable of plasminogen (PG) activation, and

XX CC fibrin binding regions of human fibrinectin, which are from fibrin

CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match 73.5%; Score 1132.6; DB 3; Length 1245;

Best Local Similarity 98.8%; Pred. No. 0;

Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 33 ATAGCTGCTCCTGAATGGCTACTAGATCGTCTCTTCTGTAAATAACAGCAATGTGTTGTT 92
DB 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 60

QY 93 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAAGCTTTAAATTTTGAATC 152

DB 61 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAAGCTTTAAATTTTGAATC 120

QY 153 GATCTAACNTCAGCACCTGCTCATGGAGGAGGACAGAGCAAGCTTAAGTCCAAATCA 212

DB 121 GATCTAACNTCAGCACCTGCTCATGGAGGAGGACAGAGCAAGCTTAAGTCCAAATCA 180

QY 213 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAAGCTGACTTACTA 272

DB 181 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAAGCTGACTTACTA 240

QY 273 AAGGCTATTCAAGAACCAATTTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGTGC 332

DB 241 AAGGCTATTCAAGAACCAATTTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGTGC 300

QY 333 ATTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTCTACTTTGCTGAC 392

DB 301 ATTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTCTACTTTGCTGAC 360

QY 393 AAAGATGTTTCGGTAACCTTTGCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT 452

DB 361 AAAGATGTTTCGGTAACCTTTGCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT 420

QY 453 GTGCGCGTTAGACCATATATAAGAAAAACCAATATACAAAAACCAAGCGAAATCTGTTGATGTG 512

DB 421 GTGCGCGTTAGACCATATATAAGAAAAACCAATATACAAAAACCAAGCGAAATCTGTTGATGTG 480

QY 513 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCAAA 572

DB 481 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCAAA 540

QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 632

DB 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 600

QY 633 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTATGAACGCTGAC 692

DB 601 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTATGAACGCTGAC 660

QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 752

DB 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 720

QY 753 TTTTACTACCGTGTATAAAATTCGGGAACAAGCTTTATAGGATCAATATAAAATCTGGTCTG 812

Db 847 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCATCAATACGTT 906
 Qy 933 GATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 992
 Db 907 GATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 966
 Qy 993 TTAGACTTCAGAGATTATACGATCCCTCGTGATAGGCTAAACTCTCTACAACAATCTC 1052
 Db 967 TTAGACTTCAGAGATTATACGATCCCTCGTGATAGGCTAAACTCTCTACAACAATCTC 1026
 Qy 1053 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGGATATACGATGAC 1112
 Db 1027 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGGATATACGATGAC 1086
 Qy 1113 ACCAACCCGATCATAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAGATGCTAGTCTAC 1172
 Db 1087 ACCAACCCGATCATAACCGTTTATATGGCAAGCGACCCGAAGGAGAGATGCTAGTCTAC 1146
 Qy 1173 CATTTAGCTGGTGGT 1187
 Db 1147 CATTTAGCTATGAT 1161

RESULT 12
 ID ABA05547
 AC ABA05547
 XX ABA05547

26-FEB-2002 (first entry)

Maxadilan-streptokinase fusion protein plasmid pTVB3maxstk.

Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
 vasodilator; thrombolytic; angina; myocardial infarction; stroke;
 gene therapy; maxadilan; sand fly; plasmid; ds.

Lutzomyia longipalpis.

Unidentified.

Synthetic.

Chimeric.

WO200185100-A2.

15-NOV-2001.

10-MAY-2001; 2001WO-US015209.

11-MAY-2000; 2000US-00569920.

(GEO) GEN HOSPITAL CORP.

Reddy VB, Lerner E;

WPI; 2002-062184/08.

New fusion protein or conjugate, useful for treating unstable angina,
 acute myocardial infarction or stroke, comprises a vasodilator
 polypeptide and a thrombolytic polypeptide, or active fragments of the
 polypeptides.

Example 1; Fig 3; 37pp; English.

The invention relates to a fusion protein or a conjugate comprising a
 vasodilator polypeptide, or its active fragment, and a thrombolytic
 polypeptide or its active fragment. The protein is useful for treating a
 subject suffering from a partially or totally occluded blood vessel,
 causing unstable angina, acute myocardial infarction or stroke. The
 polynucleotide encoding the polypeptide is useful in gene therapy. The
 polynucleotide action of the protein allows for the use of lower doses of a
 thrombolytic while maintaining the clot dissolving effectiveness of the
 thrombolytic, and the use of lower doses of the thrombolytic reduces

CC associated side effects. The present sequence is the plasmid used for the
 preparation of Maxadilan-Streptokinase fusion protein. It contains
 CC maxadilan cDNA, which encodes a vasodilator peptide produced by the
 CC salivary gland of the New World sand fly, and streptokinase cDNA
 XX
 SQ Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match 73.5%; Score 1132.6; DB 6; Length 8893;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTTCTGTTAAATAACAGCAATGTTGTTT 92
 Db 5935 ATTCTGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATGTTGTTT 5994
 Qy 93 AGCGTCTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTTTGAATC 152
 Db 5995 AGCGTCTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTTTGAATC 6054
 Qy 153 GATCTAAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 212
 Db 6055 GATCTAAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 6114
 Qy 213 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTA 272
 Db 6115 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTA 6174
 Qy 273 AAGGCTATTCAAGAAACAATGATCGCTTAACGTCCACAGTACAGAGCATCTTTGAGGTC 332
 Db 6175 AAGGCTATTCAAGAAACAATGATCGCTTAACGTCCACAGTACAGAGCATCTTTGAGGTC 6234
 Qy 333 ATTGATTTTGCAGCGGATGCAACCATTTACTGATCGAAACGCGAAGCTCTACTTTGCTGAC 392
 Db 6235 ATTGATTTTGCAGCGGATGCAACCATTTACTGATCGAAACGCGAAGCTCTACTTTGCTGAC 6294
 Qy 393 AAAGATGGTTTGGTAACTTTGCGGACCCCAACCTGTCCAAGGAATTTTGTCAAGCGGACAT 452
 Db 6295 AAAGATGGTTTGGTAACTTTGCGGACCCCAACCTGTCCAAGGAATTTTGTCAAGCGGACAT 6354
 Qy 453 GTGCGGCTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 512
 Db 6355 GTGCGGCTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 6414
 Qy 513 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACGAGTCTCAA 572
 Db 6415 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACGAGTCTCAA 6474
 Qy 573 GATACTAAGCTATTGAAAAACAACACTAGCTTATCGGTGACACCATCATCTCTCAAGAAATTA 632
 Db 6475 GATACTAAGCTATTGAAAAACAACACTAGCTTATCGGTGACACCATCATCTCTCAAGAAATTA 6534
 Qy 633 GCTCAAGCAAAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
 Db 6535 GCTCAAGCAAAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 6594
 Qy 693 TCCTCAATCGTCACTCATGACAAATGACATTTTTCGTAGGATTTTACCAATGGATCAAGAG 752
 Db 6595 TCCTCAATCGTCACTCATGACAAATGACATTTTTCGTAGGATTTTACCAATGGATCAAGAG 6654
 Qy 753 TTTTACTTTACCGTGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 6655 TTTTACTTTACCGTGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 6714
 Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGGG 872
 Db 6715 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAAGGG 6774
 Qy 873 GAAAGCCGATGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAATAGCTT 932
 Db 6775 GAAAGCCGATGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAATAGCTT 6834
 Qy 933 GATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGGACGTAAC 992

Db 6835 GATGTCGATACCAACGAATTGCTAAAGAGTAGCAGCTCTTAACAGCTAGCAACGTAAC 6894
 QY 993 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACCTACTCTACACAATCTC 1052
 Db 6895 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACCTACTCTACACAATCTC 6954
 QY 1053 GATGCTTTGGTATTATGACATATACCTTAACCTGGAAAAAGTAGAGGATATACCGATGAC 1112
 Db 6955 GATGCTTTGGTATTATGACATATACCTTAACCTGGAAAAAGTAGAGGATATACCGATGAC 7014
 QY 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGAGAGAAATGCTAGCTAC 1172
 Db 7015 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGAGAGAAATGCTAGCTAT 7074
 QY 1173 CATTTAGCTGGTGGT 1187
 Db 7075 CATTTAGCTGTATGAT 7089

RESULT 13
 AAX16632
 ID AAX16632 standard; DNA; 1242 BP.
 AC AAX16632;
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 XX Streptococcus equisimilis native streptokinase encoding DNA.
 DE Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance; ds.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1242
 FT /*tag= a
 FT /transl_except= (pos:40..42,aa:Asn)
 FT /note= "no stop codon given"
 XX
 FN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR P-PSDB; AAW94664.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 XX
 PS Claim 1; Col 7-10; 17pp; English.
 XX
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence encodes native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPlg) to
 CC plasmin (HPlm), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant

CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX Sequence 1242 BP; 423 A; 268 C; 237 G; 314 T; 0 U; 0 Other;
 SQ
 Query Match 73.4%; Score 1131; DB 2; Length 1242;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGCTCTTCTGTAAATACACGCCAATGGTTGTT 92
 Db 1 ATTGCTGGACCTCAGTGGCTGCTAGACCGTCCATCTGTACCAACGCCAATAGTTGTT 60
 QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAGACATTTAGTCTTAAATTTTTTGAATC 152
 Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAGACATTTAGTCTTAAATTTTTTGAATC 120
 QY 153 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 212
 Db 121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 180
 QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGCTCATATAAACTTGAGAAAGCTGACTTACTA 272
 Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGCTCATATAAACTTGAGAAAGCTGACTTACTA 240
 QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAAACGACGACTACTTTGAGGTC 332
 Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAAACGACGACTACTTTGAGGTC 300
 QY 333 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGCAAGGCTTAACTTTGCTGAC 392
 Db 301 ATTGATTTTCAAGAGCATGCAACCATTTACTGATGCGCAAGGCTTAACTTTGCTGAC 360
 QY 393 AAAGATGGTTTCGGTAAACCTTTGCGGACCCCAACCTGTCCAAGAAATTTTTGCTAAGCGGACAT 452
 Db 361 AAAGATGGTTTCGGTAAACCTTTGCGGACCCCAACCTGTCCAAGAAATTTTTGCTAAGCGGACAT 420
 QY 453 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512
 Db 421 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
 QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGCCAGGCTCTCAA 572
 Db 481 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGCCAGGCTCTCAA 540
 QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 632
 Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 600
 QY 633 GCTCAAGCAACAAAGCATTTTAAACCAAAACCCAGGCTATACGATTATGAACGTTGAC 692
 Db 601 GCTCAAGCAACAAAGCATTTTAAACCAAAACCCAGGCTATACGATTATGAACGTTGAC 660
 QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
 Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
 QY 753 TTTTACTACCGTGTAAAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 721 TTTTACTACCGTGTAAAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
 QY 813 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAAAATATTACGCTCTTAAAAAGGG 872
 Db 781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAAAATATTACGCTCTTAAAAAGGG 840
 QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 932
 Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 900
 QY 933 GATGTGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCAGCACTAAC 992
 Db 901 GATGTGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCAGCACTAAC 960


```

QY 993 TTAGACTTCAGAGATTATACGATCCCTCGTGATAGGCTAAACTACTCTACAACAATCTC 1052
Db |||||
QY 961 TTAGACTTCAGAGATTATACGATCCCTCGTGATAGGCTAAACTACTCTACAACAATCTC 1020
Db |||||
QY 1053 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAGGATAATCAGATGAC 1112
Db |||||
QY 1021 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAGGATAATCAGATGAC 1080
QY 1113 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAGAAATGCTAGCTAC 1172
Db |||||
QY 1081 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1140
QY 1173 CATTTAGCTGGTGGT 1187
Db |||||
QY 1141 CATTTAGCTTATGAT 1155

RESULT 14
AA16633
ID AA16633 standard; DNA; 1242 BP.
AC AA16633;
XX
DT 04-MAY-1999 (first entry)
DE Streptococcus equisimilis mutant streptokinase K59E encoding DNA.
KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance; ds.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX Synthetic.
XX
FH Key
FT CDS
FT 1..1242
FT /tag= a
FT /transl_except= (pos:40..42,aa:Aen)
FT /note= "no stop codon given"
XX
PN US5876999-A..
XX
XX 02-MAR-1999.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX (NASC-) NAT SCI COUNCIL.
XX
XX Wu H;
XX
XX WPI; 1999-189643/16.
XX P-PSDB; AA194665.
XX
XX Mutant streptokinase polypeptide - useful as plasmin-resistant
XX thrombolytic agent.
XX
XX Claim 4; Col 11-14; 17pp; English.
XX
XX The present invention describes a mutant streptokinase (SK) polypeptide
XX in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
XX segment of the corresponding native SK is replaced by another amino acid.
XX The present sequence encodes mutant SK K59E. SK is a secretory protein of
XX haemolytic Streptococcus able to activate human plasminogen (HPIg) to
XX plasmin (HPIrm), which is a serine protease able to catalyse the
XX hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
XX agent in the treatment of vascular thromboembolytic symptoms such as
XX acute myocardial infarction. Compared with wild-type SK, the K59E mutant
XX is more resistant to degradation by human plasmin and is more effective

```

```

CC both in acting as a fibrolytic agent and in activating human plasminogen
XX
SQ Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;
Query Match 73.2%; Score 1127.8; DB 2; Length 1242;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1138; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 33 ATAGCTGTCCTCGAATGCTACTAGATCGTCTCTCTCTGTGTAAATATAACAGCCCAATGGTCTT 92
Db 1 ATTGCTGACCTGATGGTCTGCTAGACCGTCACTCTGTCACCAACAGCAATAGTATGTTT 60
QY 93 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACATTAGTCTTTAAATTTTTTGAATC 152
Db 61 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACATTAGTCTTTAAATTTTTTGAATC 120
QY 153 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 212
Db 121 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCGAAATCA 180
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGGAGAAAGCTGACTTACTA 272
Db 181 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGGAGAAAGCTGACTTACTA 240
QY 273 AAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGATTAACGACGACTACTTTGAGTGC 332
Db 241 AAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGATTAACGACGACTACTTTGAGTGC 300
QY 333 ATTGATTTTGCACCGATGCAACCATTTACTGATCGAATCGGCAAGGCTTACTTTGCTGAC 392
Db 301 ATTGATTTTGCACCGATGCAACCATTTACTGATCGAATCGGCAAGGCTTACTTTGCTGAC 360
QY 393 AAAGATGGTTCGGTAACTTTGCCGACCAACCTGTCAGAAATTTTGTGAAGCGGACAT 452
Db 361 AAAGATGGTTCGGTAACTTTGCCGACCAACCTGTCAGAAATTTTGTGAAGCGGACAT 420
QY 453 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512
Db 421 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
QY 513 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 572
Db 481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 540
QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATACATCTCAAGAAATTA 632
Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATACATCTCAAGAAATTA 600
QY 633 GCTCAAGCAGAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGACGTCGAC 692
Db 601 GCTCAAGCAGAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAAACGTCGAC 660
QY 693 TCCTCAATCTGCTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAGATGGATCAAGAG 752
Db 661 TCCTCAATCTGCTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAGATGGATCAAGAG 720
QY 753 TTTACTTACCGTGTAAAAAATCGGGAAACAAGCTTTATAGGATCAATAAAAAAATCTGGTCTG 812
Db 721 TTTACTTACCGTGTAAAAAATCGGGAAACAAGCTTTATAGGATCAATAAAAAAATCTGGTCTG 780
QY 813 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 872
Db 781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 840
QY 873 GAAAGCCGATGATCCCTTTTGTATCGGAGTCACCTTGAACCTGTTTACCACATCAATAGCTT 932
Db 841 GAAAGCCGATGATCCCTTTTGTATCGGAGTCACCTTGAACCTGTTTACCACATCAATAGCTT 900
QY 933 GATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAAGCTAGCGAACGTAAC 992
Db 901 GATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAAGCTAGCGAACGTAAC 960
QY 993 TTAGACTTCAGAGATTATACGATCCCTCGTGATAGGCTAAACTACTCTACAACAATCTC 1052

```

Db	961	TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACATCTC	1020
Qy	1053	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAAAGTAGAGGATATACGATGAC	1112
Db	1021	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAAAGTAGAGGATATACGATGAC	1080
Qy	1113	ACCAACCGTATCATAAACGGTTTATATGGGCAGGCCGCCAAGGAGAGAGATGCTAGCTAC	1172
Db	1081	ACCAACCGTATCATAAACGGTTTATATGGGCAGGCCGCCAAGGAGAGAGATGCTAGCTAT	1140
Qy	1173	CATTAGCTGGTGGT	1187
Db	1141	CATTAGCCTATGAT	1155

RESULT 15	
ADMO1294	
ID	ADMO1294 standard; DNA; 7057 BP.
XX	
AC	ADMO1294;
XX	
DT	01-JUL-2004 (first entry)
XX	
DE	Plasmid pVTPA-Estrep DNA sequence.
XX	
KW	vector; plasmid; artificial intergene region; plant;
KW	transplastomic angiosperm; agronomic property; stress resistance;
KW	rbcL gene; ds.
XX	
OS	Unidentified.
XX	
PN	WO2004029256-A2.
XX	
PD	08-APR-2004.
XX	
PF	15-SEP-2003; 2003WO-CU000009.
XX	
PR	27-SEP-2002; 2002CU-00000208.
XX	
PA	(INGG-) CENT ING GENETICA & BIOTECNOLOGIA.
XX	
PI	Selman-Housein Sosa G, Aguiar Cabeza E, Gonzalez Quintero ADC;
PI	Ramos Gonzalez O;
XX	
DR	WPI; 2004-316131/29.
XX	
PT	DNA vector for transformation and expression in plastids, useful e.g. for
PT	producing pharmaceutical proteins or improving agronomic properties, has
PT	gene inserted in artificial intergene region.
XX	
PS	Example 10; SEQ ID NO 26; 74pp; Spanish.
XX	
CC	The invention relates to a DNA vector (A) for stable transformation and
CC	expression of genes (I) in plastids, where (I) is inserted in an
CC	artificial intergene region (AIR) formed by combining two 5'-untranslated
CC	regions (5'-UTRs) of genes that transcribe in different directions and
CC	are derived from plants of different divisions or classes. (A) are used
CC	to produce transplastomic angiosperms that have improved agronomic
CC	properties (e.g. resistance to biotic or abiotic stress) or express a
CC	very wide range of agricultural, veterinary, pharmaceutical, nutritional
CC	or industrial products, e.g. enzymes, vaccinating antigens, cytokines or
CC	immunoglobulins. Use of (A) eliminates the need for a transposon for gene
CC	insertion; inserted genes do not require promoters and terminators; and
CC	the structure of flanking sequences in (A) ensures universal
CC	applicability. Also any selection marker in (A) can be eliminated by
CC	homologous recombination. (A) provides efficient and stable expression of
CC	genes without causing any functional alterations. This sequence
CC	represents the plasmid pVTPA-Estrep, an example of the vector of the
CC	invention.
XX	
SQ	Sequence 7057 BP; 2029 A; 1543 C; 1608 G; 1877 T; 0 U; 0 Other;

Query Match	72.7%	Score 1120.2	DB 12	Length 7057
Best Local Similarity	97.6%	Pred. No. 2.2e-312		
Matches 1137	Conservative	0	Mismatches 28	Indels 0
QY	23	ATATACCATGATAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTCTGTCTTAATTAACAGCCA	82	
DB	2013	ATAAACCATGATTGCTTGGAGCTGAGTGGCTGCTAGACCGCTGCATCTGTCTCAACACAGCCA	2072	
QY	83	ATTGGTGTGTAGCGTGTGCTGTACTGTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATT	142	
DB	2073	ATTAGTGTGTAGCGTGTGCTGGTACTGTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATT	2132	
QY	143	TTTTGAAATCGATCTAAATCAATCAAGACCTGCTCATGGAGGAAAGACAGCAAGCTTTAAG	202	
DB	2133	TTTTGAAATTTGACCTTAACATCAAGACCTGCTCATGGAGGAAAGACAGCAAGCTTTAAG	2192	
QY	203	TCCAAATCAAAACCAATTTGCTACTGATAGTGTGGCGGATGTCATATAACTTGAGAAAGC	262	
DB	2193	TCCAAATCAAAACCAATTTGCTACTGATAGTGTGGCGGATGTCATATAACTTGAGAAAGC	2252	
QY	263	TGACTTACTAAAGGCTATTTCACAGAACAATTTGATCGCTTAAACGTCACACAGTAACGACGACTA	322	
DB	2253	TGACTTACTAAAGGCTATTTCACAGAACAATTTGATCGCTTAAACGTCACACAGTAACGACGACTA	2312	
QY	323	CTTTGAGGTCAATTGATTTTTCAGACGATGCAACCAATTACTGATCGAAACGGCAAGGTCCTA	382	
DB	2313	CTTTGAGGTCAATTGATTTTTCAGACGATGCAACCAATTACTGATCGAAACGGCAAGGTCCTA	2372	
QY	383	CTTTGCTGCAAAAGATGTTTCGGTAACTTTGCCGACCCAAACCTGTCTCAAGAAATTTTGTCT	442	
DB	2373	CTTTGCTGCAAAAGATGTTTCGGTAACTTTGCCGACCCAAACCTGTCTCAAGAAATTTTGTCT	2432	
QY	443	AACGGACATGTGCGGTTAGACATATAAGAAAGAAACCAATACAAACCAAGCGAAATC	502	
DB	2433	AACGGACATGTGCGGTTAGACATATAAGAAAGAAACCAATACAAACCAAGCGAAATC	2492	
QY	503	TGTTTGATGTGGAATATACTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACC	562	
DB	2493	TGTTTGATGTGGAATATACTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACC	2552	
QY	563	AGTCTCAAAAGATPACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCTCA	622	
DB	2553	AGTCTCAAAAGATPACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCTCA	2612	
QY	623	AGATTTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCACCGGCTATACGATTTA	682	
DB	2613	AGATTTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCACCGGCTATACGATTTA	2672	
QY	683	TGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAAT	742	
DB	2673	TGAACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAAT	2732	
QY	743	GGATCAAGAGTTTACTTTACCGTGTTTAAAAATCGGGAAACAAAGCTTTATAGGATCAATAAAAA	802	
DB	2733	GGATCAAGAGTTTACTTTACCGTGTTTAAAAATCGGGAAACAAAGCTTTATAGGATCAATAAAAA	2792	
QY	803	ATCTGCTCTGAATGAAGAAATAACACACTGACCTGATCTCTCAGAGAAATATTACGTCCT	862	
DB	2793	ATCTGCTCTGAATGAAGAAATAACACACTGACCTGATCTCTCAGAGAAATATTACGTCCT	2852	
QY	863	TAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTTACCAT	922	
DB	2853	TAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTTACCAT	2912	
QY	923	CAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAG	982	
DB	2913	CAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAG	2972	
QY	983	CGAACCTTACTTAGCTTACAGATTTTATACGATCCTCGTGATTAAGGCTTAACTACTCTA	1042	
DB	2973	CGAACCTTACTTAGCTTACAGATTTTATACGATCCTCGTGATTAAGGCTTAACTACTCTA	3032	
QY	1043	CAACAAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTTGGAAAAAGTAGAGGATAA	1102	

Db	3033	CAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGAAAGTAGAGATAA	3092
Qy	1103	TCACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGCCCGAAGGAGAGAA	1162
Db	3093	TCACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGCCCGAAGGAGAGAA	3152
Qy	1163	TGCTAGCTACCATTTAGCTGGTGGT	1187
Db	3153	TGCTAGCTATCATTTAGCCTATGAT	3177

Search completed: January 31, 2006, 18:17:49
Job time : 934.565 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:24:19 ; Search time 272.069 Seconds
(without alignments)
10068.109 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 1541
Sequence: 1 ttgtttaacttaagaagg.....ccttcacgatgttgtag 1541

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 889780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUS_COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1134	73.6	2385	3 US-09-211-542A-1	Sequence 1, Appli
2	1132.6	73.5	1242	2 US-09-211-542A-5	Sequence 5, Appli
3	1131	73.4	1242	2 US-08-568-393B-1	Sequence 1, Appli
4	1127.8	73.2	1242	2 US-08-568-393B-2	Sequence 2, Appli
5	1118.6	72.6	2566	2 US-08-488-940-19	Sequence 19, Appl
6	1111.8	72.1	1245	2 US-07-703-778D-1	Sequence 1, Appli
7	1111.6	72.1	1257	2 US-07-854-596B-25	Sequence 25, Appl
8	1110.2	72.0	1512	2 US-07-854-596B-27	Sequence 27, Appl
9	1108.6	71.9	1317	2 US-07-854-596B-18	Sequence 18, Appl
10	1108.6	71.9	1335	2 US-07-854-596B-14	Sequence 14, Appl
11	1108.6	71.9	1458	2 US-07-854-596B-42	Sequence 42, Appl
12	1108.6	71.9	1467	2 US-07-854-596B-46	Sequence 46, Appl
13	1108.6	71.9	2589	2 US-07-854-596B-34	Sequence 34, Appl
14	1084.4	70.4	1209	3 US-09-374-038-11	Sequence 11, Appl
15	1084.4	70.4	1209	3 US-09-658-179-11	Sequence 11, Appl
16	1084.4	70.4	1245	3 US-09-374-038-12	Sequence 12, Appl
17	1084.4	70.4	1245	3 US-09-658-179-12	Sequence 12, Appl
18	1083.8	70.3	1122	3 US-09-374-038-13	Sequence 13, Appl
19	1083.8	70.3	1122	3 US-09-658-179-13	Sequence 13, Appl
20	1083.8	70.3	1158	3 US-09-374-038-14	Sequence 14, Appl
21	1083.8	70.3	1158	3 US-09-658-179-14	Sequence 14, Appl
22	1078.8	70.0	2253	2 US-07-854-596B-39	Sequence 39, Appl
23	1077.6	69.9	1119	2 US-07-854-596B-30	Sequence 30, Appl
24	973	63.1	1068	3 US-09-211-542A-11	Sequence 11, Appl

25	970	62.9	2208	3 US-09-211-542A-3	Sequence 3, Appli
26	761.2	49.4	1262	9 5240845-3	Patent No. 5240845
27	756.2	49.1	1242	9 5240845-2	Patent No. 5240845
28	450	29.2	450	3 US-09-211-542A-13	Sequence 13, Appl
29	330	21.4	7679	3 US-09-220-132-38	Sequence 38, Appl
30	330	21.4	7680	3 US-09-023-855-1289	Sequence 1289, Ap
31	330	21.4	7680	6 FCT-US95-09819-6	Sequence 6, Appli
32	330	21.4	7803	2 US-08-551-356-1	Sequence 1, Appli
33	330	21.4	7803	6 FCT-US93-12687-1	Sequence 1, Appli
34	330	21.4	8044	3 US-09-566-921-135	Sequence 135, App
35	326.8	21.2	7705	2 US-08-259-569-16	Sequence 16, Appl
36	326.8	21.2	7705	2 US-08-826-885-16	Sequence 16, Appl
37	326.8	21.2	7705	9 5455158-2	Patent No. 5455158
38	50.8	3.3	986	2 US-07-637-250A-8	Sequence 8, Appli
39	50.8	3.3	986	2 US-08-145-061-8	Sequence 8, Appli
40	49.8	3.2	7218	2 US-08-232-463-14	Sequence 14, Appl
41	45.4	2.9	55	9 5240845-38	Patent No. 5240845
42	45	2.9	45	9 5455158-17	Patent No. 5455158
43	45	2.9	45	9 5455158-18	Patent No. 5455158
44	43.4	2.8	53	9 5240845-33	Patent No. 5240845
45	41.8	2.7	53	9 5240845-34	Patent No. 5240845

ALIGNMENTS

RESULT 1
US-09-211-542A-1
; Sequence 1, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2385
US-09-211-542A-1

Query Match 73.6%; Score 1134; DB 3; Length 2385;


```

QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 272
Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
QY 273 AAGGCTATTCAAGAAACAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 241 AAGGCTATTCAAGAAACAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGGTC 300
QY 333 ATTGATTTTTCGACGATGCAACCAATTTACTGATCGMAACGGCAAGGCTACTTTGCTGAC 392
Db 301 ATTGATTTTTCGACGATGCAACCAATTTACTGATCGMAACGGCAAGGCTACTTTGCTGAC 360
QY 393 AAGATGTTTTCGTAACCTTTGCGGACCCAACTGTCCAAGAAATTTTTCGTAAGCGGACAT 452
Db 361 AAGATGTTTTCGTAACCTTTGCGGACCCAACTGTCCAAGAAATTTTTCGTAAGCGGACAT 420
QY 453 GTGCGGTTAGACCATATAAAGAAACCAATTAACCAACCAAGGAAATCTGTTGATGTG 512
Db 421 GTGCGGTTAGACCATATAAAGAAACCAATTAACCAACCAAGGAAATCTGTTGATGTG 480
QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540
QY 573 GATCTAAGCTATTGAACCACTAGCTATCGGTGACACCACTCATCTCAAGAAATTA 632
Db 541 GATCTAAGCTATTGAACCACTAGCTATCGGTGACACCACTCATCTCAAGAAATTA 600
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 752
Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 720
QY 753 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 721 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 813 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
Db 781 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCATCAAAATAGCTT 932
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCATCAAAATAGCTT 900
QY 933 GATGCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 992
Db 901 GATGCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 960
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAAACAATCTC 1052
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAAACAATCTC 1020
QY 1053 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1112
Db 1021 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1080
QY 1113 ACCAACCGTATCATACCGTTTATATATGGGACGAGCCGAGAGAGATGCTAGCTAC 1172
Db 1081 ACCAACCGTATCATACCGTTTATATATGGGACGAGCCGAGAGAGATGCTAGCTAC 1140
QY 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTGGTGGT 1155

```

RESULT 3

US-08-568-393B-1

; Sequence 1, Application US/08568393B

; Patent No. 5876999

```

;
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis H46A
; INDIVIDUAL ISOLATE: Malke, H., Roe, B., and Ferretti, J. J.;
; INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A" from Gene 34:357-362 (1985).
; CELL TYPE: Streptococcus equisimilis H46A
; US-08-568-393B-1

```

```

Query Match 73.4%; Score 1131; DB 2; Length 1242;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 33 ATAGCTGCTGCTGAATGGCTACTAGATCGTCTTCTGTAATAATACAGCCAAATGCTGTTGT 92
Db 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACCAACAGCCAAATGTTGTT 60
QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
QY 153 GATCTAATCATCAGCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAATCA 212
Db 121 GATCTAATCATCAGCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAATCA 180
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 272
Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
QY 273 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 332
Db 241 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 300
QY 333 ATTGATTTTTCGACGATGCAACCAATTTACTGATCGMAACGGCAAGGCTACTTTGCTGAC 392

```

```

Db 301 ATTGATTTTGAAGCGATGCAACCAATTAATGATCGAAACGGCAAGGCTACTTTGCTGAC 360
Qy 393 AAAGATGTTTGGTAACTTTCGCGACCCCAACCTGTCGAAGAATTTTTCGTAAGCGGACAT 452
Db 361 AAAGATGTTTGGTAACTTTCGCGACCCCAACCTGTCGAAGAATTTTTCGTAAGCGGACAT 420
Qy 453 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 421 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
Qy 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGCATTTGAGACAGGCTCAAA 572
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGCATTTGAGACAGGCTCAAA 540
Qy 573 GATACTAGCTATTGAAACACCTAGCTATCGGTGACACCATCACATCTCAAGAATTAATA 632
Db 541 GATACTAGCTATTGAAACACCTAGCTATCGGTGACACCATCACATCTCAAGAATTAATA 600
Qy 633 GCTCAAGCACAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
Db 601 GCTCAAGCACAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660
Qy 693 TCCTCAATCGTCATGACCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
Db 661 TCCTCAATCGTCATGACCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
Qy 753 TTTACTTACCGTGTAAATAACCGGACCAAGCTTATAGGATCAATAAATAATCTGCTCG 812
Db 721 TTTACTTACCGTGTAAATAACCGGACCAAGCTTATAGGATCAATAAATAATCTGCTCG 780
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAGGG 872
Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTCTTAAAAAGGG 840
Qy 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCAAAATCGTT 932
Db 841 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCAAAATCGTT 900
Qy 933 GATGTCGATCAACGAATGCTAAAGTAGGAGCTCTTAAACAGCTAGCGAATGATAC 992
Db 901 GATGTCGATCAACGAATGCTAAAGTAGGAGCTCTTAAACAGCTAGCGAATGATAC 960
Qy 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACCAATCTC 1052
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACCAATCTC 1020
Qy 1053 GATGCTTTGGTATTATGAGCTATATACCTTTAACTGGAAAGTAGAGGATTAATCAGGATGAC 1112
Db 1021 GATGCTTTGGTATTATGAGCTATATACCTTTAACTGGAAAGTAGAGGATTAATCAGGATGAC 1080
Qy 1113 ACCAAGCGTATCAACCGTTTATATGCGCAAGCGACCGAAGGAGAGATGCTAGCTAC 1172
Db 1081 ACCAAGCGTATCAACCGTTTATATGCGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1140
Qy 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTGGTGGT 1155

```

RESULT 4
 US-08-568-393B-2
 ; Sequence 2, Application US/08568393B
 ; Patent No. 5876999
 ; GENERAL INFORMATION:
 ; APPLICANT: Hua-Lin Wu
 ; APPLICANT: Guey-Yueh Shi
 ; TITLE OF INVENTION: Preparation of novel streptokinase
 ; TITLE OF INVENTION: mutants as improved thrombolytic agents
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Jeing & Chang
 ; STREET: Two No. 5876999th Second Street, Suite 290
 ; CITY: San Jose

```

; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHEetical: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: SK-K59E
; LOCATION: DNA sequence No. 5876999174 and 175 have been changed
; LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
; LOCATION: from Lys to Glu.
; OTHER INFORMATION:
; US-08-568-393B-2

```

```

Query Match 73.2%; Score 1127.8; DB 2; Length 1242;
Best Local Similarity 98.5%; Pred. No. 8.7e-315; Indels 0; Gaps 0;
Matches 1138; Conservative 0; Mismatches 17;
Qy 33 ATAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTCTGTAATAACACGCAATTTGTTGT 92
Db 1 ATTGCTGGACCTTGAGTGGCTGTAGACCGTCCATCTGTACCACACGCAATTTAGTTGT 60
Qy 93 AGCGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
Db 61 AGCGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
Qy 153 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGACGAGCTTAAGTCCAAATCA 212
Db 121 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGACGAGCTTAAGTCCGGAATCA 180
Qy 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATCTGAGAAAGCTGACTTACTA 272
Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATCTGAGAAAGCTGACTTACTA 240
Qy 273 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAACGACGCTACTTTAGGTC 332
Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAACGACGCTACTTTAGGTC 300
Qy 333 ATTGATTTTGAAGCGATGCAACCATTAAGTATGATCGAAACGGCAAGGCTACTTTGCTGAC 392
Db 301 ATTGATTTTGAAGCGATGCAACCATTAAGTATGATCGAAACGGCAAGGCTACTTTGCTGAC 360
Qy 393 AAGATGTTTCGTAACTTTCCGACCCCAACCTGTCGAAGAATTTTTCGTAAGCGGACAT 452
Db 361 AAGATGTTTCGTAACTTTCCGACCCCAACCTGTCGAAGAATTTTTCGTAAGCGGACAT 420
Qy 453 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 421 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480

```


QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540
QY 573 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTAATA 632
Db 541 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTAATA 600
QY 633 GCTCAAGCACAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 601 GCTCAAGCACAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 693 TCCTCAATCGTCTCATGACCAATGACATTTTCGTAGCAATTTTACCAATGATCAAGAG 752
Db 661 TCCTCAATCGTCTCATGACCAATGACATTTTCGTAGCAATTTTACCAATGATCAAGAG 720
QY 753 TTTACTTACCGTGTAAATTCGGACACGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 721 TTTACTTACCGTGTAAATTCGGACACGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 813 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 872
Db 781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 840
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTCAACATCAATACGTT 932
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTCAACATCAATACGTT 900
QY 933 GATGTCGATACCAACGAATGCTTAAAGTGAAGAGCTCTTACAGCTAGCGAACGTAAC 992
Db 901 GATGTCGATACCAACGAATGCTTAAAGTGAAGAGCTCTTACAGCTAGCGAACGTAAC 960
QY 993 TTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAATCTCTTCAACAATCTC 1052
Db 961 TTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAATCTCTTCAACAATCTC 1020
QY 1053 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1112
Db 1021 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1080
QY 1113 ACCAACCGTATCATACCGTTTATATGCGCAACGACCCGAAAGAGAGATGCTAGCTAC 1172
Db 1081 ACCAACCGTATCATACCGTTTATATGCGCAACGACCCGAAAGAGAGATGCTAGCTAT 1140
QY 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTGGTGGT 1155

RESULT 5
US-08-488-940-19
; Sequence 19, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-488-940-19

Query Match 72.6%; Score 1118.6; DB 2; Length 2566;
Best Local Similarity 98.6%; Pred. No. 6.2e-312;
Matches 1139; Conservative 0; Mismatches 14; Indels 2; Gaps 1;
QY 33 ATAGCTGTCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGTTGTTT 92
Db 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAATGTTGTTT 956
QY 93 AGCGTTGCTGTACTGTTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTCGAATC 152
Db 957 AGCGTTGCTGTACTGTTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTCGAATC 1016
QY 153 GATCTAACATCAACACCTGCTCATGAGAGAAAGACAGACGAGGCTTAAGTCCAAATCA 212
Db 1017 GATCTAACATCAACACCTGCTCAT--AGGAAAGACAGACGAGGCTTAAGTCCAAATCA 1074
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAACTTGAGAAAGCTGACTACTA 272
Db 1075 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAACTTGAGAAAGCTGACTACTA 1134
QY 273 AAGGCTATTCAGAACAAATTTGATCGCTTAAACCTGACAGTACAGGACTACTTTGAGGTC 332
Db 1135 AAGGCTATTCAGAACAAATTTGATCGCTTAAACCTGACAGTACAGGACTACTTTGAGGTC 1194
QY 333 ATTGATTTTGCACGAGTGCACCAATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 392
Db 1195 ATTGATTTTGCACGAGTGCACCAATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 1254
QY 393 AAAGATGGTTGGTAACTTTGCGGACCCCAACTGTCGCAAGAAATTTTGTCTAAGCGGACAT 452
Db 1255 AAAGATGGTTGGTAACTTTGCGGACCCCAACTGTCGCAAGAAATTTTGTCTAAGCGGACAT 1314
QY 453 GTGGCGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 1315 GTGGCGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1374
QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 572
Db 1375 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1434
QY 573 GATACTAAGCTATTGAAAAACAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 632
Db 1435 GATACTAAGCTATTGAAAAACAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 1494
QY 633 GCTCAAGCACAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 1495 GCTCAAGCACAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTGAC 1554
QY 693 TCCTCAATCGTCTCATGACCAATGACATTTTCGTAGGATTTTACCAATGATCAAGAG 752
Db 1555 TCCTCAATCGTCTCATGACCAATGACATTTTCGTAGGATTTTACCAATGATCAAGAG 1614
QY 753 TTTACTTTACCGTGTAAAAATTCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 1615 TTTACTTTACCGTGTAAAAATTCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1674

Db 901 GATGTCACCAACGAATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
Qy 993 TTAGACTTCAGAGATTATACGATCTCGTGATAAGGCTAAACTACTCTTACAACAATCTC 1052
Db 961 TTAGACTTCAGAGATTATACGATCTCGTGATAAGGCTAAACTACTCTTACAACAATCTC 1020
Qy 1053 GATGCTTTTGGTATTATAGACTATACCTTAACTGGAAAAGTAGAGGATATACGATGAC 1112
Db 1021 GATGCTTTTGGTATTATAGACTATACCTTAACTGGAAAAGTAGAGGATATACGATGAC 1080
Qy 1113 ACCAACCGTATCATACCGTTTATATGGCAACGAGCCGGAAGAGAGAAATGCTAGCTAC 1172
Db 1081 ACCAACCGTATCATACCGTTTATATGGCAACGAGCCGGAAGAGAGAAATGCTAGCTAT 1140
Qy 1173 CATTTAGCTGGTGGT 1187
Db 1141 CATTTAGCTATGAT 1155

RESULT 7
US-07-854-596B-25;
; Sequence 25, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1257
; OTHER INFORMATION: /note= "Methionyl-streptokinase
; OTHER INFORMATION: fusion protein"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1248
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: 4..1248

US-07-854-596B-25

Query Match
Best Local Similarity 72.1%; Score 1111.6; DB 2; Length 1257;
Matches 1129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 30 ATGATAGCTGGTCTCGTAATGGCTACTAGATCGTCTTCTGTGTAATAAAGCAGCAATGGTT 89
Db 4 ATGATTCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATAGTT 63
Qy 90 GTTAGCCGTTCTGTTAGGGGACGAAATCAAGACATTAGTCTTAAATTTTTTGA 149
Db 64 GTTAGCCGTTCTGTTAGGGGACGAAATCAAGACATTAGTCTTAAATTTTTTGA 123
Qy 150 ATCGATCTAAACATCAGCACCTGCTCATGGAGGAAAGCAGAGCAAGGCTTAAGTCCAAA 209
Db 124 ATTGACCTTAACATCAGCACCTGCTCATGGAGGAAAGCAGAGCAAGGCTTAAGTCCAAA 183
Qy 210 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTTA 269
Db 184 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTTA 243
Qy 270 CTAAGGCTATTCAAGAAACAATTGATCGCTTAACGTCCACAGTAAGGAGCTACTTTGAG 329
Db 244 CTAAGGCTATTCAAGAAACAATTGATCGCTTAACGTCCACAGTAAGGAGCTACTTTGAG 303
Qy 330 GTCATTGATTTGCAAGCGATGCAACCATTAATGATCGAAACGCAAGGCTTACTTTGCT 389
Db 304 GTCATTGATTTGCAAGCGATGCAACCATTAATGATCGAAACGCAAGGCTTACTTTGCT 363
Qy 390 GACAAAGATGGTTGGTAACTTTGCCGACCCAACTGTCCAAAGAAATTTTGTAAAGCGGA 449
Db 364 GACAAAGATGGTTGGTAACTTTGCCGACCCAACTGTCCAAAGAAATTTTGTAAAGCGGA 423
Qy 450 CATGTGCGCTTAGACCATATATAAGAAAAAACCAATACAAAATCAAGCGAAATCTGTGAT 509
Db 424 CATGTGCGCTTAGACCATATATAAGAAAAAACCAATACAAAATCAAGCGAAATCTGTGAT 483
Qy 510 GTGGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCAAGGCTC 569
Db 484 GTGGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCAAGGCTC 543
Qy 570 AAAGATACTAAGCTATTGAAAAACAATGCTATCGGTGACACCATCAATCTCAAGAAATTA 629
Db 544 AAAGATACTAAGCTATTGAAAAACAATGCTATCGGTGACACCATCAATCTCAAGAAATTA 603
Qy 630 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT 689
Db 604 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT 663
Qy 690 GACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGGATTTTACCAATGGATCAA 749
Db 664 GACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGGATTTTACCAATGGATCAA 723
Qy 750 GAGTTTACTTACCGTGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 809
Db 724 GAGTTTACTTACCGTGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 783
Qy 810 CTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGCTCTTAAAAAAA 869
Db 784 CTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGCTCTTAAAAAAA 843
Qy 870 GGGGAAAGCCGTTATGATCCCTTTGATCGGAGTCACTTGAACATGTTTCAACATCAATAC 929
Db 844 GGGGAAAGCCGTTATGATCCCTTTGATCGGAGTCACTTGAACATGTTTCAACATCAATAC 903
Qy 930 GTTAGCTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGGACGT 989
Db 904 GTTAGCTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGGACGT 963
Qy 990 AACTTAGCTTCAGAGATTTTATACGATCTCTGATAGGCTAAAGCTAACTTCTTACAACAAT 1049
Db 964 AACTTAGCTTCAGAGATTTTATACGATCTCTGATAGGCTAAAGCTAACTTCTTACAACAAT 1023

QY 1165 CTAGCTACCATTTAGCTGGT 1187
 Db 1394 CTAGCTATCATTTAGCCTATGAT 1416

RESULT 9

US-07-854-596B-18
 ; Sequence 18, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1317 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..1317
 ; OTHER INFORMATION: /note= "OmpA fused to mature
 ; OTHER INFORMATION: streptokinase gene"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4..1308
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 4..1308
 ; US-07-854-596B-18

Query Match 71.9%; Score 1108.6; DB 2; Length 1317;
 Best Local Similarity 97.5%; Pred. No. 3.5e-309;
 Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCCCTGATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGGTGT 92
 Db 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGCCAAATAGTGT 126
 QY 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
 Db 127 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 186
 QY 153 GATCTAACATCAGCACTGCTCATGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 212

Db 187 GACCTAACATCAGCACTGCTCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
 QY 213 AAACCATTTGCTACTGATAGTGGCGCGATGTACACATAAATCTTGAGAAAGCTGACTTCTA 272
 Db 247 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAGAGCTGACTTCTA 306
 QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGTCACAGTAACGAGCACTACTTTGAGGTC 332
 Db 307 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGTCACAGTAACGAGCACTACTTTGAGGTC 366
 QY 333 ATTGATTTTGAAGGATGCAACCAATTTCTGATCGAAACGCGAAGGTCTACTTTGCTGAC 392
 Db 367 ATTGATTTTGAAGGATGCAACCAATTTCTGATCGAAACGCGAAGGTCTACTTTGCTGAC 426
 QY 393 AAAGATGGTTCGGTAAACCTTCGCCGACCAACTGCTGCGAAGATTTTTCCTAAGCGGACAT 452
 Db 427 AAAGATGGTTCGGTAAACCTTCGCCGACCAACTGCTGCGAAGATTTTTCCTAAGCGGACAT 486
 QY 453 GTGCGCGTTAGACCATATAAAGAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
 Db 487 GTGCGCGTTAGACCATATAAAGAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 546
 QY 513 GAATATATCTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACCGAGGTCTCAAA 572
 Db 547 GAATATATCTGTACAGTTTACTCCCTTAAACCTTGATGAGATTTTCAGACCGAGGTCTCAAA 606
 QY 573 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACCACTACATCTCAAGAAATTACTA 632
 Db 607 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACCACTACATCTCAAGAAATTACTA 666
 QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCAAGCGCTATACGATTTTGAACGTCAC 692
 Db 667 GCTCAAGCACAAAGCATTTTAAACAAACCAACCAAGCGCTATACGATTTTGAACGTCAC 726
 QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
 Db 727 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 786
 QY 753 TTTACTTACCGTGTAAATAATCGGGAAACAAAGCTTATAGATCAATAAAAAATCTGGTCTG 812
 Db 787 TTTACTTACCATGTCAAAAAATCGGGAAACAAAGCTTATAGATCAATAAAAAATCTGGTCTG 846
 QY 813 AATGAAGAAATAAACACACACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 872
 Db 847 AATGAAGAAATAAACACACACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 906
 QY 873 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCATCAAAATACGTT 932
 Db 907 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCATCAAAATACGTT 966
 QY 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 992
 Db 967 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1026
 QY 993 TTAGACTTCAGAGATTTTATAGATCTCTGATGATGAGGCTAAACTCTCTACAAATCTC 1052
 Db 1027 TTAGACTTCAGAGATTTTATAGATCTCTGATGATGAGGCTAAACTCTCTACAAATCTC 1086
 QY 1053 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATTAATCAGATGAC 1112
 Db 1087 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATTAATCAGATGAC 1146
 QY 1113 ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAC 1172
 Db 1147 ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1206
 QY 1173 CATTTAGCTGGTGGT 1187
 Db 1207 CATTTAGCTGGTGGT 1221

RESULT 10

US-07-854-596B-14
; Sequence 14, Application US/07854596B
; Patent No. 5434073

GENERAL INFORMATION:

; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McDonnell, John J

REGISTRATION NUMBER: 26,949

REFERENCE/DOCKET NUMBER: 92,337

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 1335 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..1335

OTHER INFORMATION: /note= "Streptokinase gene from S.

OTHER INFORMATION: equisimilis"

FEATURE:

NAME/KEY: CDS

LOCATION: 7..1326

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 7..1326

US-07-854-596B-14

Query Match 71.9%; Score 1108.6; DB 2; Length 1335;
Best Local Similarity 97.5%; Pred. No. 3.5e-309; Indels 0; Gaps 0;
Matches 1126; Conservative 0; Mismatches 29;

Qy	33	ATAGCTGCTCCTGAATGGCTACTAGATCGTCTCTCTGTAATAACAGCCAAATTTGGTTGTT	92
Db	85	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTTAGTTGTT	144
Qy	93	AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTTAGCTTAAATTTTTTGAATC	152
Db	145	AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTTAGCTTAAATTTTTTGAATT	204
Qy	153	GATCTAAACATACGACCTCTCATGAGGAAAGACAGACAGCGCTTAAGTCCAAATCA	212
Db	205	GACCTAAACATACGACCTCTCATGAGGAAAGACAGACAGCGCTTAAGTCCAAATCA	264
Qy	213	AAACCAATTTGCTACTGATAGTGGCGGATGTCCATATAAACTTGAGAAAGCTGCTTACTA	272
Db	265	AAACCAATTTGCTACTGATAGTGGCGGATGTCCATATAAACTTGAGAAAGCTGCTTACTA	324

RESULT 11

US-07-854-596B-42
; Sequence 42, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G

Qy	273	AAGGCTATTCAAGAAACAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTTC	332
Db	325	AAGGCTATTCAAGAAACAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTTC	384
Qy	333	ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC	392
Db	385	ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC	444
Qy	393	AAAGATGGTTGCGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT	452
Db	445	AAAGATGGTTGCGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT	504
Qy	453	GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG	512
Db	505	GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG	564
Qy	513	GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCAGGCTCTCAA	572
Db	565	GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCAGGCTCTCAA	624
Qy	573	GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA	632
Db	625	GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA	684
Qy	633	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC	692
Db	685	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC	744
Qy	693	TCCTCAATCGTCACTCATGACAAATGTCATTTTCGTTACGATTTTACCAATGGATCAAGAG	752
Db	745	TCCTCAATCGTCACTCATGACAAATGTCATTTTCGTTACGATTTTACCAATGGATCAAGAG	804
Qy	753	TTTACTTACCGTTTAAACCTCGGAAACAAGCTTATAGGATCAATAAATACTGGTCTG	812
Db	805	TTTACTTACCGTTTAAACCTCGGAAACAAGCTTATAGGATCAATAAATACTGGTCTG	864
Qy	813	AATGAAGAAATAAACACACACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAAGGG	872
Db	865	AATGAAGAAATAAACACACACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAAGGG	924
Qy	873	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTCAACCATCAATACGTT	932
Db	925	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTCAACCATCAATACGTT	984
Qy	933	GATGTCGATACCAACGAATTTGTAAGAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAC	992
Db	985	GATGTCGATACCAACGAATTTGTAAGAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAC	1044
Qy	993	TTAGACTTCAGAGATTTATACGATCCCTCGTGAAGGCTAAACTTCTACCAACATCTC	1052
Db	1045	TTAGACTTCAGAGATTTATACGATCCCTCGTGAAGGCTAAACTTCTACCAACATCTC	1104
Qy	1053	GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC	1112
Db	1105	GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC	1164
Qy	1113	ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAC	1172
Db	1165	ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAC	1224
Qy	1173	CATTAGCTGGTGGT 1187	
Db	1225	CATTAGCTGGTGGT 1239	

; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1458 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: 1..1458
 ; OTHER INFORMATION: /note= "Hirudin-streptokinase
 ; OTHER INFORMATION: fusion linked by Factor Xa cleavable IEGR"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1449
 ; NAME/KEY: mat_peptide
 ; LOCATION: 1..1449
 ; US-07-854-596B-42

Query Match 71.9%; Score 1108.6; DB 2; Length 1458;
 Best Local Similarity 97.5%; Pred. No. 3.7e-309;
 Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGTCTTCTGTAATTAACAGCCCAATTGGTTGTT 92
 Db 208 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 267
 QY 93 AGCGTTGCTGTTACTGTGTAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 152
 Db 268 AGCGTTGCTGTTACTGTGTAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 327
 QY 153 GATCTAACATCAAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAGTCCAAATCA 212
 Db 328 GACCTAACATCAAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAGTCCAAATCA 387
 QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 272
 Db 388 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACTA 447
 QY 273 AAGGCTATTCAAGAACAAATTGATCGCTTAAGTCCACAGTAAACGACGACTTTGAGTC 332
 Db 448 AAGGCTATTCAAGAACAAATTGATCGCTTAAGTCCACAGTAAACGACGACTTTGAGTC 507
 QY 333 ATTGATTTTGAAGCGATGCAACCATTTACTGTGCGAAACGCGAGGCTACTTTGCTGAC 392
 Db 508 ATTGATTTTGAAGCGATGCAACCATTTACTGTGCGAAACGCGAGGCTACTTTGCTGAC 567

QY 393 AAAAGATGGTTCGGTAAACCTTGCCGACCCAACTGTCGAAGAAATTTTGTGAAGCGGACAT 452
 Db 568 AAAAGATGGTTCGGTAAACCTTGCCGACCCAACTGTCGAAGAAATTTTGTGAAGCGGACAT 627
 QY 453 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG 512
 Db 628 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG 687
 QY 513 GAATATAGTGTACAGTTTACTCCCTTAAACCTGTAGGAGATTTCAGACCAGGCTCTCAA 572
 Db 688 GAATATAGTGTACAGTTTACTCCCTTAAACCTGTAGGAGATTTCAGACCAGGCTCTCAA 747
 QY 573 GATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 632
 Db 748 GATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 807
 QY 633 GCTCAAGCAAAAGCATTTTAAACAAAAACCAAGGCTATACGATTTATGAAACGTGAC 692
 Db 808 GCTCAAGCAAAAGCATTTTAAACAAAAACCAAGGCTATACGATTTATGAAACGTGAC 867
 QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 752
 Db 868 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 927
 QY 753 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGTATCAATAAATAATCTGTTCTG 812
 Db 928 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGTATCAATAAATAATCTGTTCTG 987
 QY 813 AATGAAGAAATAAACAACAACATGACCTGATCTCTGAGAAATATTACGCTTTAAAAAAGGG 872
 Db 988 AATGAAGAAATAAACAACAACATGACCTGATCTCTGAGAAATATTACGCTTTAAAAAAGGG 1047
 QY 873 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 932
 Db 1048 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 1107
 QY 933 GATGTCGATACCAACGAATTCCTAAAGATGACAGCTCTTAAACGCTAGGAGACGTAAC 992
 Db 1108 GATGTCGATACCAACGAATTCCTAAAGATGACAGCTCTTAAACGCTAGGAGACGTAAC 1167
 QY 993 TTAGACTTCAGAGATTTATAGATCTCTGATGAAGGCTAACTACTCTACAAATCTC 1052
 Db 1168 TTAGACTTCAGAGATTTATAGATCTCTGATGAAGGCTAACTACTCTACAAATCTC 1227
 QY 1053 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1112
 Db 1228 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1287
 QY 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAC 1172
 Db 1288 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1347
 QY 1173 CATTTAGCTGGTGGT 1187
 Db 1348 CATTTAGCTTATGAT 1362

RESULT 12
 US-07-854-596B-46
 ; Sequence 46, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL

COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1467 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1467
OTHER INFORMATION: /note= "Streptokinase-hirudin
OTHER INFORMATION: fusion linked by Factor Xa-cleavable IEGR"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1449
NAME/KEY: mat_peptide
LOCATION: 1..1449
US-07-854-596B-46

Query Match 71.9%; Score 1108.6; DB 2; Length 1467;

Best Local Similarity 97.5%; Pred. No. 3.7e-309;

Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 33 ATAGCTGGTCCGTAATGGCTACTAGATCGTCTCTGTAATAACAGCCCAATGGTGT 92
DB |||||
QY 1 ATTGTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATAGTGT 60
DB |||||
QY 93 AGCGTTGCTGTACTGTTGAGGGGACCAATCAAGACATTAGTCTTAAATTTTGAATC 152
DB |||||
QY 61 AGCGTTGCTGTACTGTTGAGGGGACCAATCAAGACATTAGTCTTAAATTTTGAAT 120
DB |||||
QY 153 GATCTAACATACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAAGTCCAAATCA 212
DB |||||
QY 121 GACCTAACATACGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAAGTCCAAATCA 180
DB |||||
QY 213 AAACATTGCTACTGATAGTGGCGGATGTCAATAAATGAGAAAGCTGACTTACTA 272
DB |||||
QY 181 AAACATTGCTACTGATAGTGGCGGATGTCAATAAATGAGAAAGCTGACTTACTA 240
DB |||||
QY 273 AAGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAACGACACTACTTTGAGGTC 332
DB |||||
QY 241 AAGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAACGACACTACTTTGAGGTC 300
DB |||||
QY 333 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 392
DB |||||
QY 301 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 360
DB |||||
QY 393 AAAGATGGTTCGGTAACTTGGCGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACAT 452
DB |||||
QY 361 AAAGATGGTTCGGTAACTTGGCGACCCAACTGTCCAGAAATTTTGTCTAAGCGGACAT 420
DB |||||
QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAAACAGCGAAATCTGTTGATGTG 512
DB |||||

DB 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
QY |||||
DB 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
DB |||||
DB 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
QY |||||
QY 573 GATACTAAGCTATTGAAAAACACATAGCTTATCGGTGACACCATCATCTCAAGAATTACTA 632
DB |||||
DB 541 GATACTAAGCTATTGAAAAACACATAGCTTATCGGTGACACCATCATCTCAAGAATTACTA 600
QY |||||
QY 633 GCTCAAGCAAAAGCATTTTAAACAAACACACCCAGGCTATACGATTTATGAACGTCGAC 692
DB |||||
DB 601 GCTCAAGCAAAAGCATTTTAAACAAACACACCCAGGCTATACGATTTATGAACGTCGAC 660
QY |||||
QY 693 TCCTCAATCGTCACTCATGACCAATGACATTTTCGTTAGGATTTTACCAATGGATCAAGAG 752
DB |||||
DB 661 TCCTCAATCGTCACTCATGACCAATGACATTTTCGTTAGGATTTTACCAATGGATCAAGAG 720
QY |||||
QY 753 TTTTACTTACCGTGTAAAAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
DB |||||
DB 721 TTTTACTTACCATGTCAAAAAATCGGAAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 780
QY |||||
QY 813 AATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 872
DB |||||
DB 781 AATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 840
QY |||||
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCACCATCAATACGTT 932
DB |||||
DB 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCACCATCAATACGTT 900
QY |||||
QY 933 GATGTGCATACCAACGAATTTGCTAAAGAGTGAGCAGCTTTAACAGCTAGCGAACGTAAC 992
DB |||||
DB 901 GATGTCAACACCAACGAATTTGCTAAAGAGTGAGCAGCTTTAACAGCTAGCGAACGTAAC 960
QY |||||
QY 993 TTAGACTTCAGAGATTTTATACGATCTCTGCTGATAAGGCTAAACTACTCTCAACAATCTC 1052
DB |||||
DB 961 TTAGACTTCAGAGATTTTATACGATCTCTGCTGATAAGGCTAAACTACTCTCAACAATCTC 1020
QY |||||
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATATACAGATGAC 1112
DB |||||
DB 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAAGATAATACAGATGAC 1080
QY |||||
QY 1113 ACCAAGCGTATACACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAC 1172
DB |||||
DB 1081 ACCAAGCGTATACACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAC 1140
QY |||||
DB 1173 CATTTAGCTGGTGT 1187
DB |||||
DB 1141 CATTTAGCTGATGAT 1155

RESULT 13

US-07-854-596B-34

; Sequence 34, Application US/07854596B

; Patent No. 5434073

; GENERAL INFORMATION:

; APPLICANT: Dawson, Keith M

; APPLICANT: Hunter, Michael G

; APPLICANT: Czaplewski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. John J. McDonnell

; STREET: Ten South Wacker Drive, Suite 3000

; CITY: Chicago

; STATE: IL

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2589 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..2589
 ; OTHER INFORMATION: /note=
 ; OTHER INFORMATION: "OmpA-Streptokinase-fusion linked
 ; OTHER INFORMATION: by thrombin-cleavable VELQGVPRG"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4..2580
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 4..2580
 ;
 ; US-07-854-596B-34

Query Match 71.9%; Score 1108.6; DB 2; Length 2589;
 Best Local Similarity 97.5%; Pred. No. 4.7e-309;
 Matches 1126; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 33 ATAGCTGGTCTGAATGGCTTACTAGATCGTCTTCTGTAATAACAGCCAAATGGTTGTT 92
 DB 67 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAGCCAAATAGTTGTT 126
 QY 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
 DB 127 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 186
 QY 153 GATCTAACTACAGCCTGCTCATGGAGGAAAGACAGCAAGCTTAACTCCAAATCA 212
 DB 187 GACCTAACTACAGCCTGCTCATGGAGGAAAGACAGCAAGCTTAACTCCAAATCA 246
 QY 213 AAACCATTTGCTACTGATAGTGGCGGATGCTACATATAACTTCAGAAAGCTGACTTACTA 272
 DB 247 AAACCATTTGCTACTGATAGTGGCGGATGCTACATATAACTTCAGAAAGCTGACTTACTA 306
 QY 273 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAACGACGACTTCTTGAGGTC 332
 DB 307 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAACGACGACTTCTTGAGGTC 366
 QY 333 ATTGATTTTGAAGCGATCAACCAATTAAGTTCGAAACGGCAAGGCTCTACTTTGCTGAC 392
 DB 367 ATTGATTTTGAAGCGATCAACCAATTAAGTTCGAAACGGCAAGGCTCTACTTTGCTGAC 426
 QY 393 AAGATGTTTGGTAACTTGGGACCAACCTGTCAGAAATTTTGTAAAGCGGACAT 452
 DB 427 AAGATGTTTGGTAACTTGGGACCAACCTGTCAGAAATTTTGTAAAGCGGACAT 486
 QY 453 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTC 512
 DB 487 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTC 546
 QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
 DB 547 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 606

QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 632
 DB 607 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 666
 QY 633 GCTCAAGCACAAGACATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
 DB 667 GCTCAAGCACAAGACATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 726
 QY 693 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTAGATTTTACCAATGGATCAAGAG 752
 DB 727 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTAGATTTTACCAATGGATCAAGAG 786
 QY 753 TTTACTTTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 DB 787 TTTACTTTACCATGTCAAAAATCGGGAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 846
 QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
 DB 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 906
 QY 873 GAAAGCCGTATGATCCCTTTTATCGGAGTCACTTGAACCTGTTCAACCATCAATACGTT 932
 DB 907 GAAAGCCGTATGATCCCTTTTATCGGAGTCACTTGAACCTGTTCAACCATCAATACGTT 966
 QY 933 GATGTCGATACCAACGAATGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 992
 DB 967 GATGTCGATACCAACGAATGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 1026
 QY 993 TTAGACTTCAGAGATTTATACGATCTCTGAGTAAGGCTAACTTACTCTACAAATCTC 1052
 DB 1027 TTAGACTTCAGAGATTTATACGATCTCTGAGTAAGGCTAACTTACTCTACAAATCTC 1086
 QY 1053 GATGCTTTTGGTATTATGAGTACTTACTTAACTCGGAAAGCTAGAGGATTAATCAGATGAC 1112
 DB 1087 GATGCTTTTGGTATTATGAGTACTTACTTAACTCGGAAAGCTAGAGGATTAATCAGATGAC 1146
 QY 1113 ACCAACCGTATCAACCGTTTATATGGGAAAGGACCCGAAAGGAGAGAGATGCTAGCTAC 1172
 DB 1147 ACCAACCGTATCAACCGTTTATATGGGAAAGGACCCGAAAGGAGAGATGCTAGCTAT 1206
 QY 1173 CATTTAGCTGGTGGT 1187
 DB 1207 CATTTAGCTGGTGGT 1221
 RESULT 14
 US-09-374-038-11
 ; Sequence 11, Application US/09374038
 ; Patent No. 6309873
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia
 ; APPLICANT: Menendez, Alina Seraleña
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6309873
 ; CURRENT APPLICATION NUMBER: US/09/374,038
 ; CURRENT FILING DATE: 1999-08-13
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; US-09-374-038-11

Query Match 70.4%; Score 1084.4; DB 3; Length 1209;
 Best Local Similarity 98.1%; Pred. No. 3.2e-302;
 Matches 1097; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 70 TAAATAACAGCCAAATGGTTGGTTAGCGTTGCTGCTACTGATGAGGAGGAAAGACAG 129
 Db 2 TGAACAAACAGCCAAATGGTTGGTTAGCGTTGCTGCTACTGATGAGGAGGAAAGACAG 61
 QY 130 TTAGTCTTAAATTTTGGTAAATCGATCTAACATCAGACCTGCTCATGAGGAGGAAAGACAG 189
 Db 62 TTAGTCTTAAATTTTGGTAAATCGATCTAACATCAGACCTGCTCATGAGGAGGAAAGACAG 121
 QY 190 ACGAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATGAGGAGGAAAGACAG 249
 Db 122 ACGAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATGAGGAGGAAAGACAG 181
 QY 250 AACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATTTGATGCTTAAAGCTCCACA 309
 Db 182 AACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATTTGATGCTTAAAGCTCCACA 241
 QY 310 GTAACGACGACTACTTTGAGGCTCATTTGATTTGCAAGCGATGCAACATTTACTGATCGAA 369
 Db 242 GTAACGACGACTACTTTGAGGCTCATTTGATTTGCAAGCGATGCAACATTTACTGATCGAA 301
 QY 370 ACGGCAAGGCTACTTTGCTGCAAAAGATGTTTCCGTTAACTTGGCGACCAACCTGTCC 429
 Db 302 ACGGCAAGGCTACTTTGCTGCAAAAGATGTTTCCGTTAACTTGGCGACCAACCTGTCC 361
 QY 430 AAGAATTTTGTCTAAGCGGACATGCGCGTTAGACCATATAAAGAAACCAATACAAA 489
 Db 362 AAGAATTTTGTCTAAGCGGACATGCGCGTTAGACCATATAAAGAAACCAATACAAA 421
 QY 490 ACCAAGCGAAATCTGTGTGATGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG 549
 Db 422 ATCAAGCGAAATCTGTGTGATGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG 481
 QY 550 AGATTTTGACACGAGTCTCAAGATGACTAAGCTATTGAAACACTAGCTATCGGTGACA 609
 Db 482 AGATTTTGACACGAGTCTCAAGATGACTAAGCTATTGAAACACTAGCTATCGGTGACA 541
 QY 610 CCATCACATCTCAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACCAAAACCCACG 669
 Db 542 CCATCACATCTCAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACCAAAACCCACG 601
 QY 670 GCTATACGATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTA 729
 Db 602 GCTATACGATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTA 661
 QY 730 CGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAATTCGGGAAACAAGCTTATA 789
 Db 662 CGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAATTCGGGAAACAAGCTTATA 721
 QY 790 GGATCAATAAATAATCTGCTGTAATGAAGAAATTAACCACTGACCTGATCTCTGAGA 849
 Db 722 AGATCAATAAATAATCTGCTGTAATGAAGAAATTAACCACTGACCTGATCTCTGAGA 781
 QY 850 AATATTACGCTCTTAAAAAGGGGAAAAAGCCGATGATCCCTTTGATCGCAGTCACTTGA 909
 Db 782 AATATTACGCTCTTAAAAAGGGGAAAAAGCCGATGATCCCTTTGATCGCAGTCACTTGA 841
 QY 910 AACTGTTTCAACATCAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAAGATGAGCAGC 969
 Db 842 AACTGTTTCAACATCAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAAGATGAGCAGC 901
 QY 970 TCTTAAAGCTAGCGAAGCTTACTGATCTCAGAGATTTTATACGATCTCTGATGATAGG 1029
 Db 902 TCTTAAAGCTAGCGAAGCTTACTGATCTCAGAGATTTTATACGATCTCTGATGATAGG 961
 QY 1030 CTAAACTACTCTACCAACATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAA 1089
 Db 962 CTAAACTACTCTACCAACATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAA 1021
 QY 1090 AAGTAGAGGATTAATCAGATGACCAACCGGATCATACCGTTTATATGGGCAAGCGAC 1149
 Db 1022 AAGTAGAGGATTAATCAGATGACCAACCGGATCATACCGTTTATATGGGCAAGCGAC 1081

QY 1150 CCGAAGGAGAGAATGCTAGCTACCATTTAGCTGGTGGT 1187
 Db 1082 CCGAAGGAGAGAATGCTAGCTATCATTTAGCCTATGAT 1119
 RESULT 15
 US-09-658-179-11
 ; Sequence 11, Application US/09658179
 ; Patent No. 6413759
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia
 ; APPLICANT: Menendez, Alina Seralena
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6413759
 ; CURRENT APPLICATION NUMBER: US/09/658,179
 ; CURRENT FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-09-658-179-11
 Query Match 70.4%; Score 1084.4; DB 3; Length 1209;
 Best Local Similarity 98.1%; Pred. No. 3-2e-302;
 Matches 1097; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 70 TAAATAACAGCCAAATGGTTGGTTAGCGTTGCTGCTACTGATGAGGAGGAAAGACAG 129
 Db 2 TGAACAAACAGCCAAATGGTTGGTTAGCGTTGCTGCTACTGATGAGGAGGAAAGACAG 61
 QY 130 TTAGTCTTAAATTTTGGTAAATCGATCTAACATCAGACCTGCTCATGAGGAGGAAAGACAG 189
 Db 62 TTAGTCTTAAATTTTGGTAAATCGATCTAACATCAGACCTGCTCATGAGGAGGAAAGACAG 121
 QY 190 AGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATGAGGAGGAAAGACAG 249
 Db 122 AGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATGAGGAGGAAAGACAG 181
 QY 250 AACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATTTGATGCTGCTTAAAGCTCCACA 309
 Db 182 AACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAATTTGATGCTGCTTAAAGCTCCACA 241
 QY 310 GTAACGACGACTACTTTGAGGCTCATTTGATTTGCAAGCGATGCAACATTTACTGATCGAA 369
 Db 242 GTAACGACGACTACTTTGAGGCTCATTTGATTTGCAAGCGATGCAACATTTACTGATCGAA 301
 QY 370 ACGGCAAGGCTACTTTGCTGCAAAAGATGTTTCCGTTAACTTGGCGACCAACCTGTCC 429
 Db 302 ACGGCAAGGCTACTTTGCTGCAAAAGATGTTTCCGTTAACTTGGCGACCAACCTGTCC 361
 QY 430 AAGAATTTTGTCTAAGCGGACATGCGCGTTAGACCATATAAAGAAACCAATACAAA 489
 Db 362 AAGAATTTTGTCTAAGCGGACATGCGCGTTAGACCATATAAAGAAACCAATACAAA 421
 QY 490 ACCAAGCGAAATCTGTGTGATGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG 549
 Db 422 ATCAAGCGAAATCTGTGTGATGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG 481
 QY 550 AGATTTTGACACGAGTCTCAAGATGACTAAGCTATTGAAACACTAGCTATCGGTGACA 609
 Db 482 AGATTTTGACACGAGTCTCAAGATGACTAAGCTATTGAAACACTAGCTATCGGTGACA 541
 QY 610 CCATCACATCTCAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACCAAAACCCACG 669
 Db 542 CCATCACATCTCAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACCAAAACCCACG 601

```
Qy 670 GCTATACGATTATGAA CGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCGGTA 729
Db |||||
Qy 602 GCTATACGATTATGAAAGGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCGGTA 661
Db |||||
Qy 730 CGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGGAAACAAGCTTATA 789
Db |||||
Qy 662 CGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAAACAAGCTTATG 721
Db |||||
Qy 790 GGATCAATAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACTGATCTCTGAGA 849
Db |||||
Qy 722 AGATCAATAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACTGATCTCTGAGA 781
Db |||||
Qy 850 AATATTACGTCCTTAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA 909
Db |||||
Qy 782 AATATTACGTCCTTAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGA 841
Db |||||
Qy 910 AACTGTTCCACATCAAAATACGTTGATGTCGATACCAACGAATGCTAAAAAGTCAGCAGC 969
Db |||||
Qy 842 AACTGTTCCACATCAAAATACGTTGATGTCGATCAACACCAACGAATGCTAAAAAGTCAGCAGC 901
Db |||||
Qy 970 TCTTAACAGCTAGGACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGAATAGG 1029
Db |||||
Qy 902 TCTTAACAGCTAGGACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGAATAGG 961
Db |||||
Qy 1030 CTAAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGAAA 1089
Db |||||
Qy 962 CTAAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGAAA 1021
Db |||||
Qy 1090 AAGTAGAGGATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGAC 1149
Db |||||
Qy 1022 AAGTAGAGGATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGAC 1081
Db |||||
Qy 1150 CCGAAGGAGAGATGCTAGCTACCAATTTAGCTGGTGGT 1187
Db |||||
Qy 1082 CCGAAGGAGAGATGCTAGCTATCATTTAGCCTATGAT 1119
Db |||||
```

Search completed: February 1, 2006, 12:43:41
Job time : 273.069 secs

1
Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**.

Because they contain data that is confidential, the results of Pending database searches should not be left in the case .

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 17:06:30 ; Search time 1273.28 Seconds
(without alignments)
10008.068 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 1541

Sequence: 1 ttgtttaaactttaagaagg.....ccttcacgatgttgtag 1541

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA Main:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1541	100.0	1541	3	US-09-940-235-9 Sequence 9, Appli
2	1496.2	97.1	2096	3	US-09-940-235-12 Sequence 12, Appl
3	1185	76.9	1661	3	US-09-940-235-10 Sequence 10, Appl
4	1179	76.5	1327	3	US-09-940-235-6 Sequence 6, Appli
5	1166.2	75.7	1377	3	US-09-940-235-5 Sequence 5, Appli
6	1134.2	73.6	1782	3	US-09-940-235-11 Sequence 11, Appl
7	1132.6	73.5	1245	3	US-09-940-235-1 Sequence 1, Appli
8	993.4	64.5	1323	8	US-10-474-792-657 Sequence 657, App
9	330	21.4	777	3	US-09-940-235-3 Sequence 3, Appli
10	330	21.4	2127	6	US-10-210-120-49 Sequence 49, Appl
11	330	21.4	2127	9	US-10-956-157-4288 Sequence 4288, Ap
12	330	21.4	2127	9	US-10-909-035-49 Sequence 49, Appl
13	330	21.4	2443	8	US-10-741-601-70 Sequence 70, Appl
14	330	21.4	2443	8	US-10-741-600-238 Sequence 238, App
15	330	21.4	2488	8	US-10-741-601-75 Sequence 75, Appl
16	330	21.4	2488	8	US-10-741-600-244 Sequence 244, App
17	330	21.4	4295	6	US-10-144-194A-51 Sequence 51, Appl
18	330	21.4	4295	8	US-10-491-566-51 Sequence 51, Appl
19	330	21.4	6510	8	US-10-741-601-72 Sequence 72, Appl
20	330	21.4	6510	8	US-10-741-600-241 Sequence 241, App
21	330	21.4	6988	7	US-10-236-392-1 Sequence 1, Appli
22	330	21.4	7361	7	US-10-236-392-3 Sequence 3, Appli
23	330	21.4	7677	9	US-10-956-157-4995 Sequence 4995, Ap

ALIGNMENTS

RESULT 1

US-09-940-235-9
; Sequence 9, Application US/09940235
; Publication No. US20030059921A1

GENERAL INFORMATION:

APPLICANT: Kumar, Rajesh

APPLICANT: Sahni, Girish

APPLICANT: Roy, Chait

APPLICANT: Rajagopal, Kammar

APPLICANT: Nihalani, Deepak

APPLICANT: Sundaram, Vasudha

APPLICANT: Yadav, Mahavir

TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE

TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION

TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID

TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 07064-009002

CURRENT APPLICATION NUMBER: US/09/940,235

CURRENT FILING DATE: 2002-04-09

PRIOR APPLICATION NUMBER: 09/471,349

PRIOR FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: IN 3825/DEL/98

PRIOR FILING DATE: 1998-12-24

NUMBER OF SEQ ID NOS: 28

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9

LENGTH: 1541

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Hybrid cassette

US-09-940-235-9

Query Match 100.0%; Score 1541; DB 3; Length 1541;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTTAACTTTAAGAAGGAGATATACCATGATAGCTGCTCTGAATGCTACTAGATC 60

Db 1 TTGTTTAACTTTAAGAAGGAGATATACCATGATAGCTGCTCTGAATGCTACTAGATC 60

QY 61 GTCTCTTCTGTAATAACAGCAATGCTGTTGTTAGCGTGTGTTGAGGGGACGA 120

Db 61 GTCTCTTCTGTAATAACAGCAATGCTGTTGTTAGCGTGTGTTGAGGGGACGA 120

QY 121 ATCAAGACATAGCTTTAAATTTTTTGAATTCATCTAACATCAGACCTGCTCATGGAG 180

Db 768 AAACCTTTTCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 827
Qy 273 AAGGCTATTCAAGAAACAAATTGATCGCTAAAGCTCCACAGTAACGACGACTACTTTGAGTGC 332
Db 828 AAGGCTATTCAAGAACAAATTGATCGCTAAAGCTCCACAGTAACGACGACTACTTTGAGTGC 887
Qy 333 ATTGATTTTGCAGGAGTCAACCAATTACTGATCGAAACGCAAGGCTACTTTGCTGAC 392
Db 888 ATTGATTTTGCAGGAGTCAACCAATTACTGATCGAAACGCAAGGCTACTTTGCTGAC 947
Qy 393 AAGATGGTTTCGTTACCTTTCGCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
Db 948 AAGATGGTTTCGTTACCTTTCGCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 1007
Qy 453 GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 1008 GTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1067
Qy 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 572
Db 1068 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 1127
Qy 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 632
Db 1128 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 1187
Qy 633 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTCGAC 692
Db 1188 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTCGAC 1247
Qy 693 TCTCAATCGTCACTCATGACATGACATTTTCGTCAGCATTTTACCAATGGATCAAGAG 752
Db 1248 TCTCAATCGTCACTCATGACATGACATTTTCGTCAGCATTTTACCAATGGATCAAGAG 1307
Qy 753 TTTTACTACCGTTTAAATTCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 1308 TTTTACTACCGTTTAAATTCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1367
Qy 813 AATGAAGAAATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 872
Db 1368 AATGAAGAAATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1427
Qy 873 GAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAACCTGTTTACCATCAATACGTT 932
Db 1428 GAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAACCTGTTTACCATCAATACGTT 1487
Qy 933 GATGTCGATACCAACGAATTGCTTAAAGTGAAGTGAAGCTCTTAAACAGCTAGCGAACGTAAC 992
Db 1488 GATGTCGATACCAACGAATTGCTTAAAGTGAAGTGAAGCTCTTAAACAGCTAGCGAACGTAAC 1547
Qy 993 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGCTTAACTCTTCAACAATCTC 1052
Db 1548 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGCTTAACTCTTCAACAATCTC 1607
Qy 1053 GATGCTTTTGTATTTATGACTATACCTTAACTGGAAGTGAAGGATATCAAGTAC 1112
Db 1608 GATGCTTTTGTATTTATGACTATACCTTAACTGGAAGTGAAGGATATCAAGTAC 1667
Qy 1113 ACCAACCGTATCATACCGTTTATATGGAAGCAAGCCGACCAAGAGAGATGCTAGCTAC 1172
Db 1668 ACCAACCGTATCATACCGTTTATATGGAAGCAAGCCGACCAAGAGAGATGCTAGCTAC 1727
Qy 1173 CATTTAGCTGGTGGCCAGCGCAACAGATTGTACCCATAGCTGAGAGTGTGTTTGTAT 1232
Db 1728 CATTTAGCTGGTGGCCAGCGCAACAGATTGTACCCATAGCTGAGAGTGTGTTTGTAT 1787
Qy 1233 CATGCTGCTGGGACTTTCCTATGTGTCGAGAAACGTTGGGAGAGCCCTTACCAAGGCTGG 1292
Db 1788 CATGCTGCTGGGACTTTCCTATGTGTCGAGAAACGTTGGGAGAGCCCTTACCAAGGCTGG 1847
Qy 1293 ATGATGGTAGATTGTTACTTTCCTGCGGAGAGGCGAGCGCATCACTTGCACTTCTAGA 1352

Db 1848 ATGATGGTAGATTGTTACTTGCCTGGGAGAAAGCGACGCGCATCACTTGCATCTTAGA 1907
Qy 1353 AATAGATGCAACAGTACAGGACACAGGACATCTTATAGAAATTGGAGACACCTGGAGCAAG 1412
Db 1908 AATAGATGCAACAGTACAGGACACAGGACATCTTATAGAAATTGGAGACACCTGGAGCAAG 1967
Qy 1413 AAGGATATCGAGGAAACCTGCTCCAGTGCATCTGCAAGCAACGCGCGAGGAGAGTGG 1472
Db 1968 AAGGATATCGAGGAAACCTGCTCCAGTGCATCTGCAAGCAACGCGCGAGGAGAGTGG 2027
Qy 1473 AAGTGTGAGAGGACACCTCTGTGACAGACCATGAGCGGATCTGGCCCTTTCACCGAT 1532
Db 2028 AAGTGTGAGAGGACACCTCTGTGACAGACCATGAGCGGATCTGGCCCTTTCACCGAT 2087
Qy 1533 GTTCGTTAG 1541
Db 2088 GTTCGTTAG 2096

RESULT 3

US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

Query Match 76.9%; Score 1185; DB 3; Length 1661;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 1 TTTGTTTAACTTTAAGAAAGGAGATATACCATGATAGTGTGCTCTGAATGGCTACTAGATC 60
Db 151 TTTGTTTAACTTTAAGAAAGGAGATATACCATGATAGTGTGCTCTGAATGGCTACTAGATC 210
Qy 61 GTCTTCTGTAAATAACAGCCAAATGGTTTGTAGCGTTGCTGTACTGTTGAGGGGACGA 120
Db 211 GTCCATCTGTCAACAAACAGCCAAATGGTTTGTAGCGTTGCTGTACTGTTGAGGGGACGA 270
Qy 121 ATCAAGACATTAGTCTTAAATTTTTCGAAATCGATCTAAACATCAGACCTGCTCATCGAG 180
Db 271 ATCAAGACATTAGTCTTAAATTTTTCGAAATCGATCTAAACATCAGACCTGCTCATCGAG 330
Qy 181 GAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 240
Db 331 GAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 390
Qy 241 TGTCAATAAACTTTGAGAAAGCTGACTTACTAAAGGCTATTTCAGAAACAATTTGATCGCTA 300

[illegible]

RESULT 4

RESUME #
US-09-940-235-6

US-03-340-233-8
; Sequence 6, Application US/09940235

; Publication No. US20030059921A1

; GENERAL INFORMATION:

```

; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-6

Query Match          76.5%; Score 1179; DB 3; Length 1327;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1182; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

RESULT 4

RESUMI 4
US-09-940-235-6

US-03-340-233-8
; Sequence 6, Application US/09940235

; Publication No. US20030059921A1

; GENERAL INFORMATION:

```
QY 661 ACCACCCAGGCTATACGATTTATGAACGAGTACTCTCAATCGTCACTCATGACAATGACA 720
Db |||||
QY 711 ACCACCCAGGCTATACGATTTATGAACGAGTACTCTCAATCGTCACTCATGACAATGACA 770
Db |||||
QY 721 TTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 780
Db |||||
QY 771 TTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 830
Db |||||
QY 781 AAGCTTATAGGATCAATAAAAAATCTGGCTGATGAAGAAATAAACAACACATGACCTGA 840
Db |||||
QY 831 AAGCTTATAGGATCAATAAAAAATCTGGCTGATGAAGAAATAAACAACACATGACCTGA 890
Db |||||
QY 841 TCTCTGAGAAATATTACGCTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
Db |||||
QY 891 TCTCTGAGAAATATTACGCTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 950
Db |||||
QY 901 GTCACCTTGAAACGTGTTACCATCAAAATAGCTTGATGTCGATACCAACGAATTCGTAATA 960
Db |||||
QY 951 GTCACCTTGAAACGTGTTACCATCAAAATAGCTTGATGTCGATACCAACGAATTCGTAATA 1010
Db |||||
QY 961 GTGAGCAGCTCTTAACAGCTAGCGAAGCTTAACCTTAGACTTCAGAGATTTATACGATCCTC 1020
Db |||||
QY 1011 GTGAGCAGCTCTTAACAGCTAGCGAAGCTTAACCTTAGACTTCAGAGATTTATACGATCCTC 1070
Db |||||
QY 1021 GTGATAAGGCTAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGCACTATACCT 1080
Db |||||
QY 1071 GTGATAAGGCTAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGCACTATACCT 1130
Db |||||
QY 1081 TAACTGGAAGATAGAGGATAATACGATGACACCAACCGTATCATACCGTTTATATGG 1140
Db |||||
QY 1131 TAACTGGAAGATAGAGGATAATACGATGACACCAACCGTATCATACCGTTTATATGG 1190
Db |||||
QY 1141 GCAAGGACCCGAGAGAGATGCTAGCTACCAATTAGCTGGTGGT 1187
Db |||||
QY 1191 GCAAGGACCCGAGAGAGATGCTAGCTATCATTTAGCCTATGAT 1237
Db |||||
```

RESULT 5

US-09-940-235-5

```
; Sequence 5, Application US/09940235
; Publication NO. US2003059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5
```

Query Match 75.7%; Score 1166.2; DB 3; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1174; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
QY 1 TTTGTTTAACTTTAAGGAAGAGATATACCATGATAGTGGTCTGTAATGGCTACTAGATC 60
Db |||||
QY 101 TTTGTTTAACTTTAAGGAAGAGAGATATACCATGATAGTGGTCTGTAATGGCTACTAGATC 160
Db |||||
QY 61 GTCCTTCTGTAATAACAGCCAAATGGTGTAGGTTGCTGGTACTGTTGAGGAGGACGA 120
Db |||||
QY 161 GTCCTATCTGTCAACAGCCAAATGGTGTAGGTTGCTGGTACTGTTGAGGAGGACGA 220
Db |||||
QY 121 ATCAAGACATTTAGTCTTAAATTTTGAATTCGATCTAACAATCAGCATCGCTCTCATGGAG 180
Db |||||
QY 221 ATCAAGACATTTAGTCTTAAATTTTGAATTCGATCTAACAATCAGCATCGCTCTCATGGAG 280
Db |||||
QY 181 GAAAGACAGAGCAAGGCTTAAAGTCCAAAATCAAAACCAATTTTGTACTGATAGTGGCGGCA 240
Db |||||
QY 281 GAAAGACAGAGCAAGGCTTAAAGTCCAAAATCAAAACCAATTTTGTACTGATAGTGGCGGCA 340
Db |||||
QY 241 TGTCAATATAACTTTGAGAAAGCTGACTTAAAGAGGCTATTCAAGAGCAATTTGATCGCTA 300
Db |||||
QY 341 TGTCAATATAACTTTGAGAAAGCTGACTTAAAGAGGCTATTCAAGAGCAATTTGATCGCTA 400
Db |||||
QY 301 ACGTCCACAGTAACGACGACTACTTTGAGGTCATTTGATTTTGAAGCGATCGAACCACTTA 360
Db |||||
QY 401 ACGTCCACAGTAACGACGACTACTTTGAGGTCATTTGATTTTGAAGCGATCGAACCACTTA 460
Db |||||
QY 361 CTGATCGAAAACGGCAAGGCTTACTTTGCTGACAAAGATGGTTGCGTAACTTTGCGGACCC 420
Db |||||
QY 461 CTGATCGAAAACGGCAAGGCTTACTTTGCTGACAAAGATGGTTGCGTAACTTTGCGGACCC 520
Db |||||
QY 421 AACCTGTCCAAAGAAATTTTGTAAAGCGGACATGTGCGGTTAGACCATATATAAGAAAAAC 480
Db |||||
QY 521 AACCTGTCCAAAGAAATTTTGTAAAGCGGACATGTGCGGTTAGACCATATATAAGAAAAAC 580
Db |||||
QY 481 CAATACAAAACCAAGCGAATCTGTTGATGTGGAAATATCTGTACAGTTACTTACTCCCTTAA 540
Db |||||
QY 581 CAATACAAAACCAAGCGAATCTGTTGATGTGGAAATATCTGTACAGTTACTTACTCCCTTAA 640
Db |||||
QY 541 ACCCTGATGACGATTTTCAAGACCGGCTCTCAAGAGATCTTAAGCTATTGAAACACATAGCTA 600
Db |||||
QY 641 ACCCTGATGACGATTTTCAAGACCGGCTCTCAAGAGATCTTAAGCTATTGAAACACATAGCTA 700
Db |||||
QY 601 TCGGTGACACCATCACAATCTCAAGAAATTAAGCTCAAGCAACAAAGCAATTTTAAACAAAA 660
Db |||||
QY 701 TCGGTGACACCATCACAATCTCAAGAAATTAAGCTCAAGCAACAAAGCAATTTTAAACAAAA 760
Db |||||
QY 661 ACCACCCAGGCTATACGATTTATGAACGAGTACTCTCAATCGTCACTCATGACAATGACA 720
Db |||||
QY 761 ACCACCCAGGCTATACGATTTATGAACGAGTACTCTCAATCGTCACTCATGACAATGACA 820
Db |||||
QY 721 TTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 780
Db |||||
QY 821 TTTTCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC 880
Db |||||
QY 781 AAGCTTATAGGATCAATAAAAAATCTGGCTGATGAAGAAATAAACAACATGACCTGA 840
Db |||||
QY 881 AAGCTTATAGGATCAATAAAAAATCTGGCTGATGAAGAAATAAACAACATGACCTGA 940
Db |||||
QY 841 TCTCTGAGAAATATTACGCTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 900
Db |||||
QY 941 TCTCTGAGAAATATTACGCTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA 1000
Db |||||
QY 901 GTCACCTTGAAACGTGTTACCATCAAAATAGCTTGATGTCGATACCAACGAATTCGTAATA 960
Db |||||
QY 1001 GTCACCTTGAAACGTGTTACCATCAAAATAGCTTGATGTCGATACCAACGAATTCGTAATA 1060
Db |||||
QY 961 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTTAACCTTAGACTTCAGAGATTTATACGATCCTC 1020
Db |||||
QY 1061 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTTAACCTTAGACTTCAGAGATTTATACGATCCTC 1120
Db |||||
QY 1021 GTGATAAGGCTAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGCACTATACCT 1080
Db |||||
QY 1121 GTGATAAGGCTAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGCACTATACCT 1180
Db |||||
QY 1081 TAACTGGAAGATAGAGGATAATACGATGACACCAACCGTATCATACCGTTTATATGG 1140
Db |||||
```

Db 1181 TAACTGAAAGTAGAGTAATACGATGACACCAACCGTATCATAAACCGTTATATGG 1240
 Qy 1141 GCAAGCGCCGAGAGAGAGTCTAGCTACCAATTTAGCTGGT 1187
 Db 1241 GCAAGCGCCGAGAGAGAGTCTAGCTATCATTTAGCCTATGAT 1287

RESULT 6

US-09-940-235-11
 ; Sequence 11, Application US/09940235
 ; Publication NO. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 1782
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hybrid cassette
 US-09-940-235-11

Query Match 73.6%; Score 1134.2; DB 3; Length 1782;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Qy 33 ATAGCTGGTCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATTTGGTTGTT 92
 Db 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTTGGTTGTT 597
 Qy 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
 Db 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 657
 Qy 153 GATCTAAATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAACTGCAAAATCA 212
 Db 658 GATCTAAATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGCTTAACTGCAAAATCA 717
 Qy 213 AAACCAATTTGCTAGTATGATGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 272
 Db 718 AAACCAATTTGCTAGTATGATGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 777
 Qy 273 AAGGCTATTCAAGAACAAATTTGATCGCTAAGCTCCAGTAAACGACGACTTTAGGTC 332
 Db 778 AAGGCTATTCAAGAACAAATTTGATCGCTAAGCTCCAGTAAACGACGACTTTAGGTC 837
 Qy 333 ATTGATTTTGAAGCGGATCAACATTAATGATCGAAACGCAAGCTTACTTTGCTGAC 392
 Db 838 ATTGATTTTGAAGCGGATCAACATTAATGATCGAAACGCAAGCTTACTTTGCTGAC 897
 Qy 393 AAAGATGGTTCGGTAACTTTCGCGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 452
 Db 898 AAAGATGGTTCGGTAACTTTCGCGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 957

RESULT 7

US-09-940-235-1
 ; Sequence 1, Application US/09940235
 ; Publication NO. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349

Qy 453 GTGCGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
 Db 958 GTGCGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
 Qy 513 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGATCTCAAA 572
 Db 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGATCTCAAA 1077
 Qy 573 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 632
 Db 1078 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
 Qy 633 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTAAACAGTGC 692
 Db 1138 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTTAAACAGTGC 1197
 Qy 693 TCCTCAATCGTCACTCATGACATTTTCCGTAACGATTTTACCAATGGATCAAGAG 752
 Db 1198 TCCTCAATCGTCACTCATGACATTTTCCGTAACGATTTTACCAATGGATCAAGAG 1257
 Qy 753 TTTTACTTACCGTGTAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 1258 TTTTACTTACCGTGTAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Qy 813 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
 Db 1318 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
 Qy 873 GAAAAAGCGGTATGATCCCTTTTGTATCGCAGTCACCTTGAACCTGTTTCAACATCAATACGTT 932
 Db 1378 GAAAAAGCGGTATGATCCCTTTTGTATCGCAGTCACCTTGAACCTGTTTCAACATCAATACGTT 1437
 Qy 933 GATGTCGATACCAACGAATTTGCTTAAAGTGAGGAGCTCTTAAAGCTAGGAAACGTAAC 992
 Db 1438 GATGTCGATACCAACGAATTTGCTTAAAGTGAGGAGCTCTTAAAGCTAGGAAACGTAAC 1497
 Qy 993 TTAGACTTCAGAGATTATACGATCTCTCGGTGATAAGGCTAACTTCTCAACAACTC 1052
 Db 1498 TTAGACTTCAGAGATTATACGATCTCTCGGTGATAAGGCTAACTTCTCAACAACTC 1557
 Qy 1053 GATGCTTTTGGTATTATGGAATACCTTAACTTGAAGGAGTATAGGAGTAAATCAGATGAC 1112
 Db 1558 GATGCTTTTGGTATTATGGAATACCTTAACTTGAAGGAGTATAGGAGTAAATCAGATGAC 1617
 Qy 1113 ACCAAGCTATCATACCGTTTATATGGCAAGGACCCGAGAGAGATGCTAGCTAC 1172
 Db 1618 ACCAAGCTATCATACCGTTTATATGGCAAGGACCCGAGAGAGATGCTAGCTAT 1677
 Qy 1173 CATTTAGCTGGTGGT 1187
 Db 1678 CATTTAGCTGGTGGT 1692

; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-09-940-235-1

Query Match 73.5%; Score 1132.6; DB 3; Length 1245;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 ATAGTGGTCTGGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGGTTGTT 92
Db |||||
1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTTGTT 60
Qy 93 AGCGTGTCTGTTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
Db |||||
61 AGCGTGTCTGTTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 120
Qy 153 GATCTAAATCAACGACCTCTCATGAGGAGAAACAGACGAGCGCTTAAGTCCAAATCA 212
Db |||||
121 GATCTAAATCAACGACCTCTCATGAGGAGAAACAGACGAGCGCTTAAGTCCAAATCA 180
Qy 213 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTACTA 272
Db |||||
181 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTACTA 240
Qy 273 AAGCTATTCAAGAACATTTGATCGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 332
Db |||||
241 AAGCTATTCAAGAACATTTGATCGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 300
Qy 333 ATTGATTTTGAAGGATCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 392
Db |||||
301 ATTGATTTTGAAGGATCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 360
Qy 393 AAAGATGGTTCGGTAACTTGGCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGACAT 452
Db |||||
361 AAAGATGGTTCGGTAACTTGGCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGACAT 420
Qy 453 GTGCGGTTAGACCATATAAAGAAACCAATAAACAACCAAGCGAAATCTGTTGATGTG 512
Db |||||
421 GTGCGGTTAGACCATATAAAGAAACCAATAAACAACCAAGCGAAATCTGTTGATGTG 480
Qy 513 GAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTGACAGCGTCTCAAA 572
Db |||||
481 GAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTGACAGCGTCTCAAA 540
Qy 573 GATCTAAGCTATTGAAACACTAGTATCGGTGACCAACCATCAATCTCAAGAAATCTACTA 632
Db |||||
541 GATCTAAGCTATTGAAACACTAGTATCGGTGACCAACCATCAATCTCAAGAAATCTACTA 600
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCGGCTATACGATTTTATGAACGTGAC 692
Db |||||
601 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCGGCTATACGATTTTATGAACGTGAC 660
Qy 693 TCCTCAATCTGTCATGACAAATGATTTTCCGTAACGATTTTACCAATGGAATCAAGAG 752
Db |||||
661 TCCTCAATCTGTCATGACAAATGATTTTCCGTAACGATTTTACCAATGGAATCAAGAG 720
Qy 753 TTTTACTTACCGTGTAAATTCGGGACCAAGCTTATAGATCAATAAATAATCTGCTCTG 812
Db |||||
721 TTTTACTTACCGTGTAAATTCGGGACCAAGCTTATAGATCAATAAATAATCTGCTCTG 780
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 872
Db |||||
781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 840

RESULT 8

US-10-474-792-657
; Sequence 657, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Laurie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 657
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-10-474-792-657

Query Match 64.5%; Score 993.4; DB 8; Length 1323;
Best Local Similarity 91.3%; Pred. No. 6.4e-275;
Matches 1054; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 33 ATAGTGGTCTGAAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGGTTGTT 92
Db |||||
79 ATTGCTGGGTATGGATGGCTACCAGACCGTCCACCTATCAATAACAGCCAGTTAGTTGTT 138
Qy 93 AGCGTGTCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db |||||
139 AGTATGCGCGGTATCGTTGAAGGTACCGATAAAAAAGTTTATATAAATTTTGAATC 198
Qy 153 GATCTAATCATCAGCACTGCTCATGAGGAGAAAGACAGACGAGCGTTAAGTCCAAATCA 212
Db |||||
199 GATCTAATCATCAGCACTGCTCATGAGGAGAAAGACAGACGAGCGTTAAGTCCAAATCA 258
Qy 213 AAACCATTTTCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 272
Db |||||
259 AAACCATTTTCTACTGATAGTGGCGCAATGCCACATAACTTGAAAGGCTGACTTATTA 318
Qy 273 AAGCTATTCAAGAACAAATTTGATCGTCAAGTCAAGTAAACGACGACTACTTTGAGGTC 332
Db |||||
319 AAAGCTATTCAAAAAACAGCTGATCGTCAAGTCAAGTAAACGACGCTACTTTGAGGTC 378
Qy 333 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 392
Db |||||

```

Db 379 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 438
Qy 393 AAAGATGTTTGGTAACCTTTCGCGCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
Db 439 AAAGATGTTTGGTAACCTTTCGCGCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 498
Qy 453 GTGCGGTTTAGACCATATAAAGAAACCAACATCAAAACCAAGCGAAATCTGTGTATGTG 512
Db 499 GTGCGGTTTAGACCATATAAAGAAACCAACATCAAAACCAAGCGAAATCTGTGTATGTG 558
Qy 513 GAATATCTGTACGTTTACTCTCTTAAACCCCTGATGAGATTTAGACAGAGTCTCAAA 572
Db 559 GAATATCTGTACGTTTACTCTCTTAAACCCCTGATGAGATTTAGACAGAGTCTCAAA 618
Qy 573 GATCTAAGCTATTCAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTAATA 632
Db 619 GATCTAAGCTATTCAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTAATA 678
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCGGCTATACGATTTTATGAACGTGAC 692
Db 679 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCGGCTATACGATTTTATGAACGTGAC 738
Qy 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTCAGTATTTTACCAATGGATCAAGAG 752
Db 739 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTCAGTATTTTACCAATGGATCAAGAG 798
Qy 753 TTTACTTACCGTGTAAATAACCGGAAACAAAGCTTATAGGATCAATTAATAAATCTGCTCG 812
Db 799 TTTACTTACCGTGTAAATAACCGGAAACAAAGCTTATAGGATCAATTAATAAATCTGCTCG 858
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAGGG 872
Db 859 AAGAAAAACGAACAACTGATCTGGTCTCTGAGAAATATTAACGCTCTTAAAAAGGG 918
Qy 873 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAGAACTGTTTCAACATCAATACGTT 932
Db 919 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAGAACTGTTTCAACATCAATACGTT 978
Qy 933 GATGTCGATACCAAGATTTGTAAGAGTGAAGGAGCTCTTAACAGCTAGCGAACGTAAC 992
Db 979 GATGTCACACCAACGAATTTGTAAGAGTGAAGGAGCTCTTAACAGCTAGCGAACGTAAC 1038
Qy 993 TTAGACTTTAGAGATTTATACATCTCTGTGATAGGCTTAACTACTCTCAACAATCTC 1052
Db 1039 TTAGACTTTAGAGATTTATACATCTCTGTGATAGGCTTAACTACTCTCAACAATCTC 1098
Qy 1053 GATGCTTTGGTATTATGAGCTATACCTTTAACTGGAAAGTAGAGGATTAATCAGATGAC 1112
Db 1099 GATGCTTTGGTATTATGAGCTATACCTTTAACTGGAAAGTAGAGGATTAATCAGGATAAG 1158
Qy 1113 ACCAACGCTATCATACCGTTTATATGGCAAGCGACCGAGGAGAGATGCTAGCTAC 1172
Db 1159 AATAATCGTGTGTTACAGTTTATATGGCAAGCGCTTAAAGGGGCAAGGGTACGTAT 1218
Qy 1173 CATTTAGCTGGTGGT 1187
Db 1219 CATTTAGCTTATGAT 1233

```

```

RESULT 9
US-09-940-235-3
; Sequence 3, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammaru
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION

```

```

; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(777)
; US-09-940-235-3

Query Match 21.4%; Score 330; DB 3; Length 777;
Best Local Similarity 100.0%; Pred. No. 7.2e-84;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1209 CCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTGTCGAGAAACG 1268
Db 448 CCCATAGCTGAGAAAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTGTCGAGAAACG 507
Qy 1269 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGGTATGATTTGTTCTTGGGAGAAAGGAGC 1328
Db 508 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGGTATGATTTGTTCTTGGGAGAAAGGAGC 567
Qy 1329 GAGCGCATCTTTCGACTTCTAGAAATAGATGATCAACGATCAGGACCAAGGACATCTCTAT 1388
Db 568 GAGCGCATCTTTCGACTTCTAGAAATAGATGATCAACGATCAGGACCAAGGACATCTCTAT 627
Qy 1389 AGAATTGGAGACACCTTGGAGCAAGAGATAATCGAGAAACCTCTCCAGTGCATCTGC 1448
Db 628 AGAATTGGAGACACCTTGGAGCAAGAGATAATCGAGAAACCTCTCCAGTGCATCTGC 687
Qy 1449 ACAGGCAACCGCCGAGGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGCGAGACCAATCG 1508
Db 688 ACAGGCAACCGCCGAGGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGCGAGACCAATCG 747
Qy 1509 AGCGGATCTGGCCCCCTTCCCGATGTTTCT 1538
Db 748 AGCGGATCTGGCCCCCTTCCCGATGTTTCT 777

```

```

RESULT 10
US-10-210-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubini, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-210-120-49

```


Query Match 21.4%; Score 330; DB 6; Length 2127;
Best Local Similarity 100.0%; Pred. No. 1.3e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGCTGGGAGAAACG 1268
DB CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGCTGGGAGAAACG 656

QY 1269 TGGGAGAAAGCCCTACCAAGCTGGATGATGGTAGATTGCTTCCCTGGGAGAAAGCAGC 1328
DB TGGGAGAAAGCCCTACCAAGCTGGATGATGGTAGATTGCTTCCCTGGGAGAAAGCAGC 716

QY 1329 GGACGCATCACTTGACCTCTTAGAATATAGATGCAAGCATCAGACACAAGGACATCCTAT 1388
DB GGACGCATCACTTGACCTCTTAGAATATAGATGCAAGCATCAGACACAAGGACATCCTAT 776

QY 1389 AGAATTGGAGACACCTGGAGCAGAGAGGATATCGAGGAAACCTGCTCCAGTGCACTGCG 1448
DB AGAATTGGAGACACCTGGAGCAGAGAGGATATCGAGGAAACCTGCTCCAGTGCACTGCG 836

QY 1449 ACAGGCAACGCCCGGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 1508
DB ACAGGCAACGCCCGGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 896

QY 1509 AGCGGATCTGGCCCTTACCGATGTTGCT 1538
DB AGCGGATCTGGCCCTTACCGATGTTGCT 926

RESULT 11
US-10-956-157-4288
; Sequence 4288, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4288
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-4288

Query Match 21.4%; Score 330; DB 9; Length 2127;
Best Local Similarity 100.0%; Pred. No. 1.3e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGCTGGGAGAAACG 1268
DB CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGCTGGGAGAAACG 656

QY 1269 TGGGAGAAAGCCCTACCAAGCTGGATGATGGTAGATTGCTTCCCTGGGAGAAAGCAGC 1328
DB TGGGAGAAAGCCCTACCAAGCTGGATGATGGTAGATTGCTTCCCTGGGAGAAAGCAGC 716

QY 1329 GGACGCATCACTTGACCTCTTAGAATATAGATGCAAGCATCAGACACAAGGACATCCTAT 1388
DB GGACGCATCACTTGACCTCTTAGAATATAGATGCAAGCATCAGACACAAGGACATCCTAT 776

QY 1389 AGAATTGGAGACACCTGGAGCAGAGAGGATATCGAGGAAACCTGCTCCAGTGCACTGCG 1448
DB AGAATTGGAGACACCTGGAGCAGAGAGGATATCGAGGAAACCTGCTCCAGTGCACTGCG 836

QY 1449 ACAGGCAACGCCCGGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 1508
DB ACAGGCAACGCCCGGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 896

QY 1509 AGCGGATCTGGCCCTTACCGATGTTGCT 1538
DB AGCGGATCTGGCCCTTACCGATGTTGCT 926

RESULT 12
US-10-909-035-49
; Sequence 49, Application US/10909035
; Publication No. US20050136493A1
; GENERAL INFORMATION:
; APPLICANT: Rubin, Mark A.
; APPLICANT: Chinmaiyan, Arul M.
; APPLICANT: Laxman, Bharathi
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: AMACR Cancer Markers
; FILE REFERENCE: UM-09098
; CURRENT APPLICATION NUMBER: US/10/909,035
; CURRENT FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-909-035-49

Query Match 21.4%; Score 330; DB 9; Length 2127;
Best Local Similarity 100.0%; Pred. No. 1.3e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGCTGGGAGAAACG 1268
DB CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGCTGGGAGAAACG 656

QY 1269 TGGGAGAAAGCCCTACCAAGCTGGATGATGGTAGATTGCTTCCCTGGGAGAAAGCAGC 1328
DB TGGGAGAAAGCCCTACCAAGCTGGATGATGGTAGATTGCTTCCCTGGGAGAAAGCAGC 716

QY 1329 GGACGCATCACTTGACCTCTTAGAATATAGATGCAAGCATCAGACACAAGGACATCCTAT 1388
DB GGACGCATCACTTGACCTCTTAGAATATAGATGCAAGCATCAGACACAAGGACATCCTAT 776

QY 1389 AGAATTGGAGACACCTGGAGCAGAGAGGATATCGAGGAAACCTGCTCCAGTGCACTGCG 1448
DB AGAATTGGAGACACCTGGAGCAGAGAGGATATCGAGGAAACCTGCTCCAGTGCACTGCG 836

QY 1449 ACAGGCAACGCCCGGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 1508
DB ACAGGCAACGCCCGGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCAATCG 896

QY 1509 AGCGGATCTGGCCCTTACCGATGTTGCT 1538
DB AGCGGATCTGGCCCTTACCGATGTTGCT 926

RESULT 13
US-10-741-601-70
; Sequence 70, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-70

Db 61 GTCCCTTCGTAATAACAGCCCAATGGTGGTGTAGCGTTCGTGTA CTGTTGAGGGGACGA 120
Qy 121 ATCAAGACATTAAGTCTTAAATTTTTTGAATCGATCTAAACATCAAGCACTGCTCATGGAG 180
Db 121 ATCAAGACATTAAGTCTTAAATTTTTTGAATCGATCTAAACATCAAGCACTGCTCATGGAG 180
Qy 181 GAAAGACAGAGCAAGGCTTAAGTCCAAAATCAAAACCAATTTGCTACTGTAGTGGCGCGA 240
Db 181 GAAAGACAGAGCAAGGCTTAAGTCCAAAATCAAAACCAATTTGCTACTGTAGTGGCGCGA 240
Qy 241 TGTCACTAATACTTGGAAGCTGACTTAAGGCTATTCAAGACCAATTCATCGCTA 300
Db 241 TGTCACTAATACTTGGAAGCTGACTTAAGGCTATTCAAGACCAATTCATCGCTA 300
Qy 301 ACGTCCACAGTAACGACGACTACTTTGAGGTCAATGATTTTGAAGCGATGCAACCATTA 360
Db 301 ACGTCCACAGTAACGACGACTACTTTGAGGTCAATGATTTTGAAGCGATGCAACCATTA 360
Qy 361 CTGATCGAAAACGGCAAGGTCTACTTTGCTGACAAAAGATGGTTGGTAACTTCCGACCC 420
Db 361 CTGATCGAAAACGGCAAGGTCTACTTTGCTGACAAAAGATGGTTGGTAACTTCCGACCC 420
Qy 421 AACCTGTCCAGAAATTTTGTGTAAGCGGACATGTGCGCGTTAGACCATATAAGAAAAAC 480
Db 421 AACCTGTCCAGAAATTTTGTGTAAGCGGACATGTGCGCGTTAGACCATATAAGAAAAAC 480
Qy 481 CAATACAAAACCAAGCGAATCTGTTGATGCGAATATCTGACAGTTTACTCCCTTAA 540
Db 481 CAATACAAAACCAAGCGAATCTGTTGATGCGAATATCTGACAGTTTACTCCCTTAA 540
Qy 541 ACCCTGATGACGATTTTCAGACAGGTCTCAAAAGATACTAAGCTATTGAAAAACACTAGCTA 600
Db 541 ACCCTGATGACGATTTTCAGACAGGTCTCAAAAGATACTAAGCTATTGAAAAACACTAGCTA 600
Qy 601 TCGGTGACACATCATCTCAAGAAATTAAGTCTCAAGCAAAAGCAATTTTAAACAAA 660
Db 601 TCGGTGACACATCATCTCAAGAAATTAAGTCTCAAGCAAAAGCAATTTTAAACAAA 660
Qy 661 ACCACCCAGGCTATACGATTTATGACGAGTCTCTCAATCGTCTCAATCGTCTCAATCGTCTCA 720
Db 661 ACCACCCAGGCTATACGATTTATGACGAGTCTCTCAATCGTCTCAATCGTCTCAATCGTCTCA 720
Qy 721 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTGTTAAAAATCGGGAAC 780
Db 721 TTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTGTTAAAAATCGGGAAC 780
Qy 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 840
Db 781 AAGCTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGA 840
Qy 841 TCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAGCGGTATGATCCGTTTATGTCGA 900
Db 841 TCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAGCGGTATGATCCGTTTATGTCGA 900
Qy 901 GTCACTTGAAACTGTTTCAACATCAATACGTTGATGTCATACCAACGAATGCTAAAAA 960
Db 901 GTCACTTGAAACTGTTTCAACATCAATACGTTGATGTCATACCAACGAATGCTAAAAA 960
Qy 961 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTTAAGTCTGAGACTTCAGAGATTTATACGATCCTC 1020
Db 961 GTGAGCAGCTCTTAAACAGCTAGCGAAGCTTAAGTCTGAGACTTCAGAGATTTATACGATCCTC 1020
Qy 1021 GTGATAAGGCTTAACTACTCTCAACATCTCGATGCTTTTGGTATTATGACTATACCT 1080
Db 1021 GTGATAAGGCTTAACTACTCTCAACATCTCGATGCTTTTGGTATTATGACTATACCT 1080
Qy 1081 TAACTGGAAGAGTAGAGGATATCAGATGACCAACCGGTATCATACCGTTTATATGG 1140
Db 1081 TAACTGGAAGAGTAGAGGATATCAGATGACCAACCGGTATCATACCGTTTATATGG 1140
Qy 1141 GCAAGCGACCCGAGAGAGAAATGCTAGCTACCATTTAGCTGTGTGGCGAGCGCAAC 1200
Db 1141 GCAAGCGACCCGAGAGAGAAATGCTAGCTACCATTTAGCTGTGTGGCGAGCGCAAC 1200

Qy 1201 AGATTGTACCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCTTATGTGTCG 1260
Db 1201 AGATTGTACCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCTTATGTGTCG 1260
Qy 1261 GAGAAACGTGGGAGAGCCCTTACCAAGCTGGATGATGGTAGATTGTACTTGCCTGGGAG 1320
Db 1261 GAGAAACGTGGGAGAGCCCTTACCAAGCTGGATGATGGTAGATTGTACTTGCCTGGGAG 1320
Qy 1321 AAGGCAAGCGGACCCATCACTTGCATCTCTAGAAATAGATGCAAGGATCAGACACAAGGA 1380
Db 1321 AAGGCAAGCGGACCCATCACTTGCATCTCTAGAAATAGATGCAAGGATCAGACACAAGGA 1380
Qy 1381 CATCTCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGT 1440
Db 1381 CATCTCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGT 1440
Qy 1441 GCATCTGCACAGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGACACCTCTGTGCGAGA 1500
Db 1441 GCATCTGCACAGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGACACCTCTGTGCGAGA 1500
Qy 1501 CCACATCGAGCGGATCTGGGCCCTTCAACCGATGTTGGTTAG 1541
Db 1501 CCACATCGAGCGGATCTGGGCCCTTCAACCGATGTTGGTTAG 1541

RESULT 2
US-10-631-558-12
; Sequence 12, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-12

Query Match 97.1%; Score 1496.2; DB 7; Length 2096;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 33 ATAGCTGGTCTCTGAATGGCTACTAGATCGTCCTTCTGTAATAAATCAAGCAATTTGGTTGT 92
Db 588 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACCAATTTGGTTGT 647
Qy 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
Db 648 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 707

Oy 153 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAAATCA 212
 Db 708 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGGCTTAAGTCCAAAATCA 767
 Oy 213 AAACCATTTGCTACTGATGTCGCGCATGCTCATTAATTGAGAAAGCTGACTTA 272
 Db 768 AAACCATTTGCTACTGATGTCGCGCATGCTCATTAATTGAGAAAGCTGACTTA 827
 Oy 273 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGCAAGGTCTACTTGTCTGAC 332
 Db 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGCAAGGTCTACTTGTCTGAC 887
 Oy 333 ATTTGATTTTGGCAAGCGATGCAACATTTAGTGAAGCGGCAAGGTCTACTTGTCTGAC 392
 Db 888 ATTTGATTTTGGCAAGCGATGCAACATTTAGTGAAGCGGCAAGGTCTACTTGTCTGAC 947
 Oy 393 AAAAGATGCTTGGCTTAACCTTGGCCGACCCGACCTGTCCAGAAATTTTGTCTAGCGGACAT 452
 Db 948 AAAAGATGCTTGGCTTAACCTTGGCCGACCCGACCTGTCCAGAAATTTTGTCTAGCGGACAT 1007
 Oy 453 GTGGCGCTTGAACCATATTAAGAAAAACCAATACAAAGCGAAATCTGTTGATGTG 512
 Db 1008 GTGGCGCTTGAACCATATTAAGAAAAACCAATACAAAGCGAAATCTGTTGATGTG 1067
 Oy 513 GAATATCTACTGTAAGTTTACTCTCCCTTAAACCTGATGACGATTTGACAGCGGTCTCAAA 572
 Db 1068 GAATATCTACTGTAAGTTTACTCTCCCTTAAACCTGATGACGATTTGACAGCGGTCTCAAA 1127
 Oy 573 GATATCTAGCTTATGAAAAACATAGTATGCTGACACCATCATCTCAAGAAATTA 632
 Db 1128 GATATCTAGCTTATGAAAAACATAGTATGCTGACACCATCATCTCAAGAAATTA 1187
 Oy 633 GCTCAAGCAAGGCAATTTTAAACAAACCAACCGGCTATGATTTTGAACCGGAC 692
 Db 1188 GCTCAAGCAAGGCAATTTTAAACAAACCAACCGGCTATGATTTTGAACCGGAC 1247
 Oy 693 TCCCTCAATGCTACTCATGACATGATCATTTTCCGTAAGATTTTCAATGATGATCAAG 752
 Db 1248 TCCCTCAATGCTACTCATGACATGATCATTTTCCGTAAGATTTTCAATGATGATCAAG 1307
 Oy 753 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGCTG 812
 Db 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGCTG 1367
 Oy 813 AATGAAGAAATTAACAACATGACCTGATCTGGAAGAAATTAATGCTTTAAAAAGG 872
 Db 1368 AATGAAGAAATTAACAACATGACCTGATCTGGAAGAAATTAATGCTTTAAAAAGG 1427
 Oy 873 GAAAAGCCGTATGATCCCTTGTATGTCAGTCACTTGAACCTGTTCAACATCAATGCT 932
 Db 1428 GAAAAGCCGTATGATCCCTTGTATGTCAGTCACTTGAACCTGTTCAACATCAATGCT 1487
 Oy 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAACAGCTTTTAAACAGTACGGAAGTAA 992
 Db 1488 GATGTCGATACCAACGAATTTGCTAAAAAGTGAACAGCTTTTAAACAGTACGGAAGTAA 1547
 Oy 993 TTGAATCTCAGAGTTTATGATCTCTGATTAAGGCTTAACTACTTCAACAATCTC 1052
 Db 1548 TTGAATCTCAGAGTTTATGATCTCTGATTAAGGCTTAACTACTTCAACAATCTC 1607
 Oy 1053 GATGCTTTTGTGATTAATGATCTATACCTTAACTGGAAGAAATGATATATCAAGTATGAC 1112
 Db 1608 GATGCTTTTGTGATTAATGATCTATACCTTAACTGGAAGAAATGATATATCAAGTATGAC 1667
 Oy 1113 ACCAAGCGTATCATTAACGCTTTATATGAGGCAAGCGGCAAGGAGAAATGCTAGTAC 1172
 Db 1668 ACCAAGCGTATCATTAACGCTTTATATGAGGCAAGCGGCAAGGAGAAATGCTAGTAC 1727
 Oy 1173 CATTTAGTGTGTGTGGCCAGGCGCAACATTTGTAACCTATAGCTGAAAGTGTGTTGAT 1232
 Db 1728 CATTTAGTGTGTGTGGCCAGGCGCAACATTTGTAACCTATAGCTGAAAGTGTGTTGAT 1787
 Oy 1233 CATGCTGTGGGACTTCTATGTGTGGAGAAACGTGGAGAAACCTTACCAAGGCTG 1292

Db 1788 CATGCTGTGGGACTTCTATGTGTGGAGAAACGTGGAGAAACCTTACCAAGGCTG 1847
 Oy 1293 ATGATGTGATGATTTTACTTGTGCTGAGGAAAGCGAGCGGACCATCATCTTGA 1352
 Db 1848 ATGATGTGATGATTTTACTTGTGCTGAGGAAAGCGAGCGGACCATCATCTTGA 1907
 Oy 1353 AATAGATGCAACGATCAGGACCAAGGACATCTTATAGAAATTTGAGAACCTGAGCAAG 1412
 Db 1908 AATAGATGCAACGATCAGGACCAAGGACATCTTATAGAAATTTGAGAACCTGAGCAAG 1967
 Oy 1413 AAGATATATGAGGAAACCTGCTCAAGTATCTTGAACAGGCAAGCGGCGAGAGAGTGG 1472
 Db 1968 AAGATATATGAGGAAACCTGCTCAAGTATCTTGAACAGGCAAGCGGCGAGAGAGTGG 2027
 Oy 1473 AAGTGTGAGGACCACTCTGTGTGACACCAATGAGCGGATCTGGCCCTTCAAGCAT 1532
 Db 2028 AAGTGTGAGGACCACTCTGTGTGACACCAATGAGCGGATCTGGCCCTTCAAGCAT 2087
 Oy 1533 GTTCTTAG 1541
 Db 2088 GTTCTTAG 2096

RESULT 3
 US-10-631-558-10
 ; Sequence 10, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammarra
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vaesudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 1661
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hybrid cassette
 US-10-631-558-10

Query Match 76.9%; Score 1185; DB 7; Length 1661;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1 TTTGTTAACTTTAAGAGAGATATATCAATGATAGCTGTGCTGAATGCTACTAGATC 60
 Db 151 TTTGTTAACTTTAAGAGAGATATATCAATGATAGCTGTGCTGAATGCTACTAGATC 210
 Oy 61 GTCCCTCTGTAAATTAACAGCAATTTGCTTTAGCTGTGCTGTAAGGAGGAGCA 120
 Db 211 GTCCATCTGTCAACAAACAGCAATTTGCTTTAGCTGTGCTGTAAGGAGGAGCA 270
 Oy 121 ATCAAGCATTAAGCTTTAAATTTTGAATCGATCTTAATCAAGCACTGCTCATAGAG 180


```

Db      531 CATACAAAACGAGAAATCTGTTGATGTGAAATTAAGTACAGTTACTCCCTTAA 590
Qy      541 ACCCTGATGATGATTTTCAGACGAGTCTCAAGATTAAGTATTAAGAAAACTAGCTA 600
Db      591 ACCCTGATGATGATTTTCAGACGAGTCTCAAGATTAAGTATTAAGAAAACTAGCTA 650
Qy      601 TCGGTGACACATCACTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 660
Db      651 TCGGTGACACATCACTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 710
Qy      661 ACCACGAGCTTAAGTATTAAGAAAGTCAAGTCCCAATGCACTCAAGCAATGACA 720
Db      711 ACCACGAGCTTAAGTATTAAGAAAGTCAAGTCCCAATGCACTCAAGCAATGACA 770
Qy      721 TTTTCGTAAGATTTTACCAATGATCAAGATTAAGTCAAGTCCCAATGCAAGTCA 780
Db      771 TTTTCGTAAGATTTTACCAATGATCAAGATTAAGTCAAGTCCCAATGCAAGTCA 830
Qy      781 AAGCTTATAGATCAATTAATAAAATCTGTTGAAATGAAGAAATTAACACACTGACCTGA 840
Db      831 AAGCTTATAGATCAATTAATAAAATCTGTTGAAATGAAGAAATTAACACACTGACCTGA 890
Qy      841 TCTCTGAGAAATTAAGTCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db      891 TCTCTGAGAAATTAAGTCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 950
Qy      901 GTCACTTGAAGCTTTTCAACATCAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 960
Db      951 GTCACTTGAAGCTTTTCAACATCAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1010
Qy      961 GTGAGAGCTCTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1020
Db      1011 GTGAGAGCTCTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1070
Qy      1021 GTGATTAAGCTTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1080
Db      1071 GTGATTAAGCTTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1130
Qy      1081 TAACTGGAAGAAAGTGAAGTAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1140
Db      1131 TAACTGGAAGAAAGTGAAGTAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1190
Qy      1141 GCAACGAGCCGAGAGAGAAATGTAGCTAACATTTAGCTGAGT 1187
Db      1191 GCAACGAGCCGAGAGAGAAATGTAGCTAACATTTAGCTGAGT 1237

RESULT 5
US-10-631-558-5
; Sequence 5, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammaru
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEIN POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT FILING DATE: 2003-07-31
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-631-558-5

Query Match      75.7%; Score 1166.2; DB 7; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1174; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy      1 TTTGTTTAACTTTTAAAGAGATATTAACATGATAGTGTCTGAAATGCTACTAGATC 60
Db      101 TTTGTTTAACTTTTAAAGAGATATTAACATGATAGTGTCTGAAATGCTACTAGATC 160
Qy      61 GTCTCTGTAATTAACAGCCAAATGCTGTTGAAGTGTGCTGTAATGCTGTAAGGAGCA 120
Db      161 GTCTCTGTAATTAACAGCCAAATGCTGTTGAAGTGTGCTGTAATGCTGTAAGGAGCA 220
Qy      121 ATCAAGACATTAAGCTTTAAATTTTGAATATGATCTTAACATCAAGCTGCTGATGAG 180
Db      221 ATCAAGACATTAAGCTTTAAATTTTGAATATGATCTTAACATCAAGCTGCTGATGAG 280
Qy      181 GAAAGACAGAGCAAGGCTTAAGTCAAAATCAAAACATTTGCTACTGATAGTGGCGCA 240
Db      281 GAAAGACAGAGCAAGGCTTAAGTCAAAATCAAAACATTTGCTACTGATAGTGGCGCA 340
Qy      241 TGTCAATTAACCTTGAAGAAAGCTGATCTTAAGGCTATTAAGAAATTAAGTATGCTA 300
Db      341 TGTCAATTAACCTTGAAGAAAGCTGATCTTAAGGCTATTAAGAAATTAAGTATGCTA 400
Qy      301 AAGTCCACAGTAAGAGAGCACTTGAAGTCAATTTGTAAGTCAAGGAGTCAAGTCAAGTCA 360
Db      401 AAGTCCACAGTAAGAGAGCACTTGAAGTCAATTTGTAAGTCAAGGAGTCAAGTCAAGTCA 460
Qy      361 CTGATCGAAAACGAGCAAGTCTACTTGTCTGCAAAAGATGTTGCGTAACTTGGCGACCC 420
Db      461 CTGATCGAAAACGAGCAAGTCTACTTGTCTGCAAAAGATGTTGCGTAACTTGGCGACCC 520
Qy      421 AACTGTCAAGAAATTTTGTAAAGCGAGCATGTGGCGTTAAGCAATTAAGAAAAAC 480
Db      521 AACTGTCAAGAAATTTTGTAAAGCGAGCATGTGGCGTTAAGCAATTAAGAAAAAC 580
Qy      481 CAATTAAGAAACCAAGCGAAATCTGTGATGAGAAATATATCTGTAAGTTTACTCCCTTAA 540
Db      581 CAATTAAGAAACCAAGCGAAATCTGTGATGAGAAATATATCTGTAAGTTTACTCCCTTAA 640
Qy      541 ACCGTATGAGATTTTCAAGCAAGGCTCAAAAGATTAAGTCAAGTCAAGTCAAGTCAAGTCA 600
Db      641 ACCGTATGAGATTTTCAAGCAAGGCTCAAAAGATTAAGTCAAGTCAAGTCAAGTCAAGTCA 700
Qy      601 TCGGTGACACATCACTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 660
Db      701 TCGGTGACACATCACTCAAGAAATTAAGTCAAGCAAAAGATTTTAAACAAA 760
Qy      661 ACCACGAGCTTAAGTATTAAGAAAGTCAAGTCCCAATGCACTCAAGCAATGACA 720
Db      761 ACCACGAGCTTAAGTATTAAGAAAGTCAAGTCCCAATGCACTCAAGCAATGACA 820
Qy      721 TTTTCGTAAGATTTTACCAATGATCAAGATTAAGTCAAGTCCCAATGCAAGTCAAGTCA 780
Db      821 TTTTCGTAAGATTTTACCAATGATCAAGATTAAGTCAAGTCCCAATGCAAGTCAAGTCA 880
Qy      781 AAGCTTATAGATCAATTAATAAAATCTGTTGAAATGAAGAAATTAACACACTGACCTGA 840
Db      881 AAGCTTATAGATCAATTAATAAAATCTGTTGAAATGAAGAAATTAACACACTGACCTGA 940
Qy      841 TCTCTGAGAAATTAAGTCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db      941 TCTCTGAGAAATTAAGTCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1000
Qy      901 GTCACTTGAAGCTTTTCAACATCAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 960

```


Db 1001 GTCACTTGAAAGCTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTCCTAAAAA 1060
QY 961 GTGAGCAGCTCTTAACAGCTACGCAAGTAACCTTAGACTTCAGAGATTTATACGATCCCTC 1020
Db 1061 GTGAGCAGCTCTTAACAGCTACGCAAGTAACCTTAGACTTCAGAGATTTATACGATCCCTC 1120
QY 1021 GTGATAAGGCTAAACTACTCTCTACAACAATCTCGATGCTTTTGGTATTTATGGACTATACCT 1080
Db 1121 GTGATAAGGCTAAACTACTCTCTACAACAATCTCGATGCTTTTGGTATTTATGGACTATACCT 1180
QY 1081 TAACTGGAAGAATAGAGATATACGATGACCAACCGTATCATAACCGTTTATATGG 1140
Db 1181 TAACTGGAAGAATAGAGATATACGATGACCAACCGTATCATAACCGTTTATATGG 1240
QY 1141 GCAAGCGACCCGAGGAGATGCTAGCTACCAATTTAGCTGGTGGT 1187
Db 1241 GCAAGCGACCCGAGGAGATGCTAGCTACCAATTTAGCTGGTGGT 1287

RESULT 6
US-10-631-558-11
; Sequence 11, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette

US-10-631-558-11

Query Match 73.6%; Score 1134.2; DB 7; Length 1782;
Best Local Similarity 98.9%; Pred. No. 0; Mismatches 13; Indels 0; Gaps 0;
Matches 1142; Conservative 0;

QY 33 ATAGCTGCTCCTGAATGGCTACTAGATCGTCTCTCTGTAATAACAGCAATTTGGTTGT 92
Db 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGGTTGT 597
QY 93 AGCGTTGCTGTTAGTGGGACGAATCAAGACATTTAGTCTTTAAATTTTGGAAATC 152
Db 598 AGCGTTGCTGTTAGTGGGACGAATCAAGACATTTAGTCTTTAAATTTTGGAAATC 657
QY 153 GATCTAAATCAGCACTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
Db 658 GATCTAAATCAGCACTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 717
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTTGAGAAAGCTGACTACTA 272
Db 718 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTTGAGAAAGCTGACTACTA 777

RESULT 7
US-10-631-558-1
; Sequence 1, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish

QY 273 AAGGCTATTCAAGAAACAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTTGAGGTC 332
Db 778 AAGGCTATTCAAGAAACAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTTGAGGTC 837
QY 333 ATTGATTTTTCGAAGCGATGCAACCAATTTATGATCGAAACGGCAAGGCTACTTTTGGCTGAC 392
Db 838 ATTGATTTTTCGAAGCGATGCAACCAATTTATGATCGAAACGGCAAGGCTACTTTTGGCTGAC 897
QY 393 AAGATGGTTTCGTTAACCCTTGGCGACCAACCAATTTATGATCGAAACGGCAAGGCTACTTTTGGCTGAC 452
Db 898 AAGATGGTTTCGTTAACCCTTGGCGACCAACCAATTTATGATCGAAACGGCAAGGCTACTTTTGGCTGAC 957
QY 453 GTGCGGTTAGACCATATATAAGAAACCAATTAACCAACCAAGCAAAATCTGTTGATGTG 512
Db 958 GTGCGGTTAGACCATATATAAGAAACCAATTAACCAACCAAGCAAAATCTGTTGATGTG 1017
QY 513 GAATATACTGTACAGTTTACTCCCTTTAAACCCCTGTGATGACGATTTTCAGACAGGCTCTCAAA 572
Db 1018 GAATATACTGTACAGTTTACTCCCTTTAAACCCCTGTGATGACGATTTTCAGACAGGCTCTCAAA 1077
QY 573 GATACTAAGCTATGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 632
Db 1078 GATACTAAGCTATGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 1137
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGACGATGAC 692
Db 1138 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGACGATGAC 1197
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGTGATCGATTTTACCAATGGATCAAGAG 752
Db 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTGATCGATTTTACCAATGGATCAAGAG 1257
QY 753 TTTACTTACCGTGTAAAAAATCGGGAACAAAGCTTTATGATGATCAATAAAAAATCTGGTCTG 812
Db 1258 TTTACTTACCGTGTAAAAAATCGGGAACAAAGCTTTATGATGATCAATAAAAAATCTGGTCTG 1317
QY 813 AATGAAGAAATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 872
Db 1318 AATGAAGAAATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1377
QY 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACCTTTACCATCAAAATAGCTT 932
Db 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACCTTTACCATCAAAATAGCTT 1437
QY 933 GATGTCGATACCAACGAATTTGCTTAAAAAGTGACGAGCTCTTTAACAGCTAGCGAAACGTAAC 992
Db 1438 GATGTCGATACCAACGAATTTGCTTAAAAAGTGACGAGCTCTTTAACAGCTAGCGAAACGTAAC 1497
QY 993 TTAGACTTCAGAGATTTTATACGATCTCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1052
Db 1498 TTAGACTTCAGAGATTTTATACGATCTCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1557
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATCAGATGAC 1112
Db 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATCAGATGAC 1617
QY 1113 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGTAC 1172
Db 1618 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGTAC 1677
QY 1173 CATTTAGCTGGTGGT 1187
Db 1678 CATTTAGCTGGTGGT 1692

APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammara
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
PRIOR FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1245
TYPE: DNA
ORGANISM: Streptococcus equisimilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1242)
US-10-631-558-1

Query Match 73.5%; Score 1132.6; DB 7; Length 1245;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1141; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 33 ATAGTGTGCTCGATGCTGCTACTAGTCTGCTCTCTGTAATAACGCCAATTGTTGTT 92
DB 1 ATGTGTGACCTGAGGCTGCTGACCGCTCACTGTCACAAACGCAATTAGTGTGTT 60
QY 93 AGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 152
DB 61 AGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 153 GATCTACATCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 212
DB 121 GATCTACATCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 213 AAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 272
DB 181 AAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 273 AAGGCTATTCAGAACCAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 332
DB 241 AAGGCTATTCAGAACCAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 333 ATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392
DB 301 ATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 393 AAAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
DB 361 AAAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 453 GTGCGCGTTGACCACTATTAAGAAAAACCAATACCAAAACCAAGCGAAATCTGTTGATG 512
DB 421 GTGCGCGTTGACCACTATTAAGAAAAACCAATACCAAAACCAAGCGAAATCTGTTGATG 480
QY 513 GAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572
DB 481 GAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 573 GATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 632
DB 541 GATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 633 GCTCAGACCAAGACATTTTAAACAAAACCAAGGCTATAGCAATTTATGAACGTGAC 692
DB 601 GCTCAGACCAAGACATTTTAAACAAAACCAAGGCTATAGCAATTTATGAACGTGAC 660
QY 693 TCTCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752
DB 661 TCTCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 753 TTTACTACCGTGTAAATATGGAACCAAGGCTATAGCAATTTATGAACGTGAC 812
DB 721 TTTACTACCGTGTAAATATGGAACCAAGGCTATAGCAATTTATGAACGTGAC 780
QY 813 AATGAGAAATTAACCAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 872
DB 781 AATGAGAAATTAACCAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 873 GAAAAGCGTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932
DB 841 GAAAAGCGTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 933 GATGTGATCCAAAGCAATGCTTAAAGTGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 992
DB 901 GATGTGATCCAAAGCAATGCTTAAAGTGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 960
QY 993 TTGACTTCAAGATTTATAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1052
DB 961 TTGACTTCAAGATTTATAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1053 GATGCTTTGCTATTTATGACCTATACCTTAACTGAAAAGTGAAGATTAACGATGAC 1112
DB 1021 GATGCTTTGCTATTTATGACCTATACCTTAACTGAAAAGTGAAGATTAACGATGAC 1080
QY 1113 ACCAAGCGTATCAATACCGTTTATATGCGCAAGGACCCGAAAGAGAAATGCTAGTAC 1172
DB 1081 ACCAAGCGTATCAATACCGTTTATATGCGCAAGGACCCGAAAGAGAAATGCTAGTAT 1140
QY 1173 CATTAGCTGATGAT 1187
DB 1141 CATTAGCTGATGAT 1155

RESULT 8
US-10-631-558-3
; Sequence 3, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (1) ... (777)
US-10-631-558-3

Query Match 21.4%; Score 330; DB 7; Length 777;
Best Local Similarity 100.0%; Pred. No. 4.4e-95;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTTGTCGGAGAAACG 1268
DB 448 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTTGTCGGAGAAACG 507
QY 1269 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTAATTCCTCGGAGAAAGCAGC 1328
DB 508 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTAATTCCTCGGAGAAAGCAGC 567
QY 1329 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1388
DB 568 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 627
QY 1389 AGAATTGGAGACACCTGGAGCAGAGGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
DB 628 AGAATTGGAGACACCTGGAGCAGAGGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 687
QY 1449 ACAGGCAACGCCGAGGAGAGTGGAGTGTGAGAGGCACACCTCTGTGAGGACCAATCG 1508
DB 688 ACAGGCAACGCCGAGGAGAGTGGAGTGTGAGAGGCACACCTCTGTGAGGACCAATCG 747
QY 1509 AGCGGATCGGCCCTTCACCGATGTTGCT 1538
DB 748 AGCGGATCGGCCCTTCACCGATGTTGCT 777

RESULT 9

US-10-995-561-114
; Sequence 114, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-114

Query Match 21.4%; Score 330; DB 7; Length 2443;
Best Local Similarity 100.0%; Pred. No. 9.1e-95;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTTGTCGGAGAAACG 1268
DB 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTTGTCGGAGAAACG 972
QY 1269 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTAATTCCTCGGAGAAAGCAGC 1328
DB 973 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTAATTCCTCGGAGAAAGCAGC 1032
QY 1329 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1388
DB 1033 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 1389 AGAATTGGAGACACCTGGAGCAGAGGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACACCTGGAGCAGAGGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152

QY 1449 ACAGGCAACGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 1508
DB 1153 ACAGGCAACGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 1212
QY 1509 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1538
DB 1213 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1242

RESULT 10

US-10-995-561-105
; Sequence 105, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-105

Query Match 21.4%; Score 330; DB 7; Length 2488;
Best Local Similarity 100.0%; Pred. No. 9.2e-95;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTTGTCGGAGAAACG 1268
DB 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTTGTCGGAGAAACG 972
QY 1269 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTAATTCCTCGGAGAAAGCAGC 1328
DB 973 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTAATTCCTCGGAGAAAGCAGC 1032
QY 1329 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1388
DB 1033 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 1389 AGAATTGGAGACACCTGGAGCAGAGGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACACCTGGAGCAGAGGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
QY 1449 ACAGGCAACGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 1508
DB 1153 ACAGGCAACGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGAGACCACATCG 1212
QY 1509 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1538
DB 1213 AGCGGATCTGGCCCCCTTCACCGATGTTGCT 1242

RESULT 11

US-10-995-561-112
; Sequence 112, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112

LENGTH: 6510
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-112

Query Match 21.4%; Score 330; DB 7; Length 6510;
Best Local Similarity 100.0%; Pred. No. 1.7e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAGGTTTGTATCATGCTGCGGAGCTTCCTATGTTGTCGGAAGC 1268
DB 913 CCCATAGCTGAGAGGTTTGTATCATGCTGCGGAGCTTCCTATGTTGTCGGAAGC 972
QY 1269 TGGGAGAAAGCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAGGAGC 1328
DB 973 TGGGAGAAAGCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAGGAGC 1032
QY 1329 GGAAGCATCACTTGCATCTTGAATAATGATGCAAGATGAGACACAGACATCTTAT 1388
DB 1033 GGAAGCATCACTTGCATCTTGAATAATGATGCAAGATGAGACACAGACATCTTAT 1092
QY 1389 AGAATTGGAGACCTCTGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACCTCTGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC 1152
QY 1449 ACAGGCAAGCGCCGAGAGAGTGAAGTGAAGGACACCTCTGTCAGACACATCG 1508
DB 1153 ACAGGCAAGCGCCGAGAGAGTGAAGTGAAGGACACCTCTGTCAGACACATCG 1212
QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTGCT 1538
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTGCT 1242

RESULT 12
US-10-995-561-117
Sequence 117, Application US/10995561
Publication No. US20050272054A1

GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 117
LENGTH: 7823
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-117

Query Match 21.4%; Score 330; DB 7; Length 7823;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAGGTTTGTATCATGCTGCGGAGCTTCCTATGTTGTCGGAAGC 1268
DB 913 CCCATAGCTGAGAGGTTTGTATCATGCTGCGGAGCTTCCTATGTTGTCGGAAGC 972
QY 1269 TGGGAGAAAGCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAGGAGC 1328
DB 973 TGGGAGAAAGCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAGGAGC 1032
QY 1329 GGAAGCATCACTTGCATCTTGAATAATGATGCAAGATGAGACACAGACATCTTAT 1388
DB 1033 GGAAGCATCACTTGCATCTTGAATAATGATGCAAGATGAGACACAGACATCTTAT 1092
QY 1389 AGAATTGGAGACCTCTGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACCTCTGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC 1152

QY 1449 ACAGGCAAGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTCAGACACATCG 1508
DB 1153 ACAGGCAAGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTCAGACACATCG 1212
QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTGCT 1538
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTGCT 1242

RESULT 13
US-10-995-561-111
Sequence 111, Application US/10995561
Publication No. US20050272054A1

GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 111
LENGTH: 7848
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-111

Query Match 21.4%; Score 330; DB 7; Length 7848;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAGGTTTGTATCATGCTGCGGAGCTTCCTATGTTGTCGGAAGC 1268
DB 913 CCCATAGCTGAGAGGTTTGTATCATGCTGCGGAGCTTCCTATGTTGTCGGAAGC 972
QY 1269 TGGGAGAAAGCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAGGAGC 1328
DB 973 TGGGAGAAAGCCCTACCAAGGCTGATGATGATTTGTACTTCCCTGGAGAGGAGC 1032
QY 1329 GGAAGCATCACTTGCATCTTGAATAATGATGCAAGATGAGACACAGACATCTTAT 1388
DB 1033 GGAAGCATCACTTGCATCTTGAATAATGATGCAAGATGAGACACAGACATCTTAT 1092
QY 1389 AGAATTGGAGACCTCTGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACCTCTGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCTGC 1152
QY 1449 ACAGGCAAGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTCAGACACATCG 1508
DB 1153 ACAGGCAAGCGCCGAGAGAGTGAAGTGAAGGACACACCTCTGTCAGACACATCG 1212
QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTGCT 1538
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTGCT 1242

RESULT 14
US-10-995-561-113
Sequence 113, Application US/10995561
Publication No. US20050272054A1

GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 113
; LENGTH: 7935
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-113

Query Match 21.4%; Score 330; DB 7; Length 7935;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACATTCCTATGTTGGTCGAGAAACG 1268
DB 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACATTCCTATGTTGGTCGAGAAACG 972
QY 1269 TGGGAGAGCCCTACCAAGCTGGATGATGTTAGTGTGCTGGGAGAGCAGC 1328
DB 973 TGGGAGAGCCCTACCAAGCTGGATGATGTTAGTGTGCTGGGAGAGCAGC 1032
QY 1329 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1388
DB 1033 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1092
QY 1389 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
DB 1093 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
QY 1449 ACAGGCACCGCGGAGGAGTGGAGTGTGAGAGGACACCTCTGTGCAGACCATCG 1508
DB 1153 ACAGGCACCGCGGAGGAGTGGAGTGTGAGAGGACACCTCTGTGCAGACCATCG 1212
QY 1509 AGCGGATCTGGCCCTTACCGATGTTTGGT 1538
DB 1213 AGCGGATCTGGCCCTTACCGATGTTTGGT 1242

RESULT 15

US-10-995-561-108
; Sequence 108, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 7959
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-108;

Query Match 21.4%; Score 330; DB 7; Length 7959;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACATTCCTATGTTGGTCGAGAAACG 1268
DB 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACATTCCTATGTTGGTCGAGAAACG 972
QY 1269 TGGGAGAGCCCTACCAAGCTGGATGATGTTAGTGTGCTGGGAGAGCAGC 1328
DB 973 TGGGAGAGCCCTACCAAGCTGGATGATGTTAGTGTGCTGGGAGAGCAGC 1032
QY 1329 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1388
DB 1033 GGACGCATCATTGCACTTTAGAAATAGATCAACGATCAGGACACAAAGGACATCCTAT 1092
QY 1389 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448

DB 1093 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
QY 1449 ACAGGCACCGCGGAGGAGTGGAGTGTGAGAGGACACACCTCTGTGCGAGACCATCG 1508
DB 1153 ACAGGCACCGCGGAGGAGTGGAGTGTGAGAGGACACACCTCTGTGCGAGACCATCG 1212
QY 1509 AGCGGATCTGGCCCTTACCGATGTTTGGT 1538
DB 1213 AGCGGATCTGGCCCTTACCGATGTTTGGT 1242

Search completed: February 1, 2006, 14:43:12
Job time : 285.911 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 6332.25 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 1541

Sequence: 1 ttgtgttaacttaagaag.....cctcacgagatgtcgttag 1541

Scoring table: IDENTITY NUC

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:
10: gb_est10:
11: gb_est11:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	330	21.4	451	1	AL603368 DKFZp686C
2	330	21.4	560	7	CN419585 170005318
3	330	21.4	621	7	CN482442 h20d08.Y
4	330	21.4	626	7	CN419611 170005313
5	330	21.4	737	1	AU140993 AU140993
6	330	21.4	765	3	BM715855 UI-E-B70-
7	330	21.4	861	1	AU140971 AU140971
8	330	21.4	7501	4	BC078656 Homo sapi
9	330	21.4	7777	4	BC100030 Homo sapi
10	330	21.4	7777	4	CR749316 Homo sapi
11	330	21.4	7868	4	BSM80653 Homo sapi
12	330	21.4	7885	4	CR749281 Homo sapi
13	330	21.4	8411	4	BSM80692 Homo sapi
14	329	21.3	465	1	AI095589 gbd2408.x
15	329	21.3	734	1	AU140802 AU140802
16	328.4	21.3	8121	4	CR749317 Homo sapi
17	326.8	21.2	548	6	CD613789 56022208U
18	325.2	21.1	551	6	CD613788 56022208U
19	325.2	21.1	560	1	AI743013 w85a12.x
20	319	20.7	500	2	BP956982 OVA-NN114
21	318.2	20.6	547	2	BE009640 PM4-BN017
22	318	20.6	943	5	BX391752 BX391752

23	317	20.6	705	1	AU140789
24	315.4	20.5	739	1	AL706215
25	310	20.1	603	7	CN419658
26	307	19.9	414	1	AA492032
27	305	19.8	518	1	AI093548
28	304	19.7	7434	11	DQ039102
29	298	19.3	538	1	AL603599
30	297.4	19.3	496	2	BG945197
31	296.4	19.2	303	1	AA852090
32	287	18.6	427	7	CN419650
33	287	18.6	452	7	CN419649
34	281	18.2	605	2	CN419657
35	280	18.2	408	7	BG900107
36	277.8	18.0	580	3	BQ292415
37	274	17.8	495	2	BP956977
38	271.8	17.6	494	2	BP96365
39	268.6	17.4	484	2	B1058354
40	268	17.4	725	5	BQ574857
41	264	17.1	454	8	W46530
42	261.8	17.0	474	6	CB536952
43	256.8	16.7	645	8	DN123922
44	251.8	16.3	314	7	CN419631
45	249.4	16.2	899	5	BX327266

ALIGNMENTS

RESULT 1
AL603368 451 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp686C067.5, mRNA sequence.
DEFINITION DKFZp686C067.5, mRNA sequence.
ACCESSION AL603368
VERSION AL603368.1 GI:15166874
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS Ansoorge,W., Winkner,U., Mewes,W., Weil,B. and Wiemann,S.
TITLES EST (Ansoorge,W., Winkner,U., Mewes,W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories), Heidelberg/Germany) within the CDNA sequencing consortium of the German Genome Project.
No 5' sequence available.
This clone (DKFZp686C067) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcencentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
location/Qualifiers

FEATURES

source
1..451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686C067"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="686 (synonym: h1cc3)"
/note="Vector: pT7Blue2; Site_1: SfiIA; Site_2: SfiIB; CDNA-collection"

ORIGIN

Query Match 21.4%; Score 330; DB 1; Length 451;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTGGTCGAGAAACG 1268
|||||
Db 32 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTGGTCGAGAAACG 91
|||||
QY 1269 TGGGAGAAGCCCTACCAAGGCTGGATGGTATGATTGTAATCTTGGCCCTGGGAGAAGGCAGC 1328
|||||
Db 92 TGGGAGAAGCCCTACCAAGGCTGGATGGTATGATTGTAATCTTGGCCCTGGGAGAAGGCAGC 151
|||||
QY 1329 GGACGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 1388
|||||
Db 152 GGACGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 211
|||||
QY 1389 AGAATTGGAGACACCTTGGAGGAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
|||||
Db 212 AGAATTGGAGACACCTTGGAGGAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 271
|||||
QY 1449 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAACATCG 1508
|||||
Db 272 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAACATCG 331
|||||
QY 1509 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 1538
|||||
Db 332 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 361
|||||

RESULT 2
CN419585
LOCUS CN419585 560 bp mRNA linear EST 16-MAY-2004
DEFINITION 1700531856084 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN419585
VERSION CN419585.1 GI:47407179
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 560)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rac,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE Transcription characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 560 Std Error: 0.00.
FEATURES
Location/Qualifiers
1..560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN
Query Match 21.4%; Score 330; DB 7; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.2e-82;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTGGTCGAGAAACG 1268
|||||
Db 209 CCCATAGCTGAGAAGTGTGTTTGGATCATGCTGCTGGGACTTCCTATGTGGTCGAGAAACG 268
|||||
QY 1269 TGGGAGAAGCCCTACCAAGGCTGGATGGTATGATTGTAATCTTGGCCCTGGGAGAAGGCAGC 1328
|||||
Db 269 TGGGAGAAGCCCTACCAAGGCTGGATGGTATGATTGTAATCTTGGCCCTGGGAGAAGGCAGC 328
|||||
QY 1329 GGACGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 1388
|||||
Db 329 GGACGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 388
|||||
QY 1389 AGAATTGGAGACACCTTGGAGGAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 1448
|||||
Db 389 AGAATTGGAGACACCTTGGAGGAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGC 448
|||||
QY 1449 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAACATCG 1508
|||||
Db 449 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAACATCG 508
|||||
QY 1509 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 1538
|||||
Db 509 AGCGGATCTGGCCCTTCCACCGATGTTTGGT 538
|||||

RESULT 3
CN482442
LOCUS CN482442 621 bp mRNA linear EST 26-APR-2004
DEFINITION hw20d08.y1 Human primary human ocular pericytes. Unamplified (hw)
CN482442
ACCESSION Homo sapiens cDNA clone hw20d08 5', mRNA sequence.
VERSION CN482442.1 GI:46563946
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 621)
AUTHORS Tsai,J.Y. and Wistow,G.
TITLE Expressed sequence tag analysis of cultured primary human ocular
pericytes
JOURNAL Unpublished (2004)
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 20 row: d column: 08
Seq primer: M13RP1 reverse primer (ABI).
FEATURES
Location/Qualifiers
1..621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hw20d08"
/cell_type="pericytes"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human primary human ocular pericytes.
Unamplified (hw)"
/note="Organ: Eye; Vector: pSPORT1; RNA was extracted from
primary human pericytes in culture. A directionally cloned
cDNA library in the pSPORT1 vector (Invitrogen) was
constructed at Bioserve Biotechnology (Laurel MD)
essentially following the protocols of the SuperScript
Plasmid System full details of which are contained in the
manufacturer's instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adaptor
[5'-pGACTAGTCTTAGATCGGCGCGGCC(T)15-3']. cDNA was


```

Query Match      21.4%; Score 330; DB 1; Length 737;
Best Local Similarity 100.0%; Pred. No. 1.3e-82;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 1268
Db 362 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 421

QY 1269 TGGGAGAAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAGGCGAGC 1328
Db 422 TGGGAGAAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAGGCGAGC 481

QY 1329 GGAAGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 1388
Db 482 GGAAGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 541

QY 1389 AGAATGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCAGTGCATCTGC 1448
Db 542 AGAATGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCAGTGCATCTGC 601

QY 1449 ACAGGCAACGGCGAGGAGAGTGGAAAGTGTGAGAGGACACACCTCTGTGCAGACCAACATCG 1508
Db 602 ACAGGCAACGGCGAGGAGAGTGGAAAGTGTGAGAGGACACACCTCTGTGCAGACCAACATCG 661

QY 1509 AGCGGATCTGGCCCTTCACCGATGTTTTCGT 1538
Db 662 AGCGGATCTGGCCCTTCACCGATGTTTTCGT 691

RESULT 6
BM715855
LOCUS      765 bp      mRNA      linear      EST 28-FEB-2002
DEFINITION      UI-E-EJ0-abj-h-11-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone
ACCESSION      BM715855
VERSION      BM715855.1 GI:19029113
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 765)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
PUBMED      8889548
COMMENT      Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse
Location/Qualifiers
1. .765
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-abj-h-11-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
FEATURES
source

```

```

/clone lib="UI-E-EJ0"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATCAAGA; lens, CGATTAGCGA; eye anterior segment,
AATGCCGCGAT; optic nerve, CCATTAGTGT; retina, CGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."
```

ORIGIN

Query Match 21.4%; Score 330; DB 3; Length 765;
Best Local Similarity 100.0%; Pred. No. 1.3e-82;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 1268
Db 182 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 241

QY 1269 TGGGAGAAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAGGCGAGC 1328
Db 242 TGGGAGAAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTTGGGAGAAGGCGAGC 301

QY 1329 GGAAGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 1388
Db 302 GGAAGCATCACTTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 361

QY 1389 AGAATGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCAGTGCATCTGC 1448
Db 362 AGAATGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCAGTGCATCTGC 421

QY 1449 ACAGGCAACGGCGAGGAGAGTGGAAAGTGTGAGAGGACACACCTCTGTGCAGACCAACATCG 1508
Db 422 ACAGGCAACGGCGAGGAGAGTGGAAAGTGTGAGAGGACACACCTCTGTGCAGACCAACATCG 481

QY 1509 AGCGGATCTGGCCCTTCACCGATGTTTTCGT 1538
Db 482 AGCGGATCTGGCCCTTCACCGATGTTTTCGT 511

RESULT 7
AUI40971
LOCUS 861 bp mRNA linear EST 05-AUG-2002
DEFINITION PLACE4 Homo sapiens cDNA clone PLACE4000583 5', mRNA
ACCESSION AUI40971
VERSION AUI40971.1 GI:11002492
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 861)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
TITLE HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES

source
 1. 861
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE4000583"
 /issue_type="placenta"
 /clone_lib="PLACE4"
 /note="Vector: pME18SFL3"

ORIGIN

Query Match 21.4%; Score 330; DB 1; Length 861;
 Best Local Similarity 100.0%; Pred. No. 1.4e-82;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 1268
 |||||
 DB 158 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 217
 |||||
 QY 1269 TGGGAGAGCCCTACCAAGGCTGGATGATGTATGTTCTGCTGGAGAGAGCAGC 1328
 |||||
 DB 218 TGGGAGAGCCCTACCAAGGCTGGATGATGTATGTTCTGCTGGAGAGAGCAGC 277
 |||||
 QY 1339 GGAGCGATCATCTTCACTTAGAATAGATGACATCAGACACAGACATCTTAT 1388
 |||||
 DB 278 GGAGCGATCATCTTCACTTAGAATAGATGACATCAGACACAGACATCTTAT 337
 |||||
 QY 1389 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTCTCCAGTGCATCTGC 1448
 |||||
 DB 338 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTCTCCAGTGCATCTGC 397
 |||||
 QY 1449 AAGGCAAGCCCGGAGAGAGTGTGAGAGGACACCTCTGTGCGAGACACATCG 1508
 |||||
 DB 398 AAGGCAAGCCCGGAGAGAGTGTGAGAGGACACCTCTGTGCGAGACACATCG 457
 |||||
 QY 1509 AGCGGATCTGGCCCTTCAACCGATGTTCTG 1538
 |||||
 DB 458 AGCGGATCTGGCCCTTCAACCGATGTTCTG 487
 |||||

RESULT 8
 BC078656 7501 bp mRNA linear HTC 21-JUL-2005
 LOCUS BC078656
 DEFINITION Homo sapiens cDNA clone IMAGE:30347017, containing frame-shift
 errors.
 ACCESSION BC078656
 VERSION BC078656.1 GI:50925326
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 7501)

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haib, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y.K., S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kelleman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smal, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 PUBMED
 12477932
 2 (bases 1 to 7501)

REFERENCE
 AUTHORS NIH MGC Project
 DIRECT SUBMISSION
 TITLE Submitted (02-AUG-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: Dr. Stefan Hanson
 cDNA Library Preparation: Michael Brownstein / Ted Usdin
 Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@pdxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 168 Row: 0 Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 47132556
 This clone has the following problem: frame shifted.

FEATURES

source
 1. 7501
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30347017"
 /issue_type="placenta, normal"
 /clone_lib="NIH_MGC_147"
 /lab_host="DH10B"
 /note="Vector: pBluescriptR"

ORIGIN

Query Match 21.4%; Score 330; DB 4; Length 7501;
 Best Local Similarity 100.0%; Pred. No. 2.5e-82;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 1268
 |||||
 DB 806 CCCATGCTGAGAGAGTTTGTATCATCTGCTGGAGCTTCTATGTGTGGAGAAACG 865
 |||||
 QY 1269 TGGGAGAGCCCTACCAAGGCTGGATGATGTATGTTCTGCTGGAGAGAGCAGC 1328
 |||||
 DB 866 TGGGAGAGCCCTACCAAGGCTGGATGATGTATGTTCTGCTGGAGAGAGCAGC 925
 |||||
 QY 1339 GGAGCGATCATCTTCACTTAGAATAGATGACATCAGACACAGACATCTTAT 1388
 |||||
 DB 926 GGAGCGATCATCTTCACTTAGAATAGATGACATCAGACACAGACATCTTAT 985
 |||||
 QY 1389 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTCTCCAGTGCATCTGC 1448
 |||||
 DB 986 AGAATTGGAGACACTCTGAGCAAGAGATATCGAGAAACCTCTCCAGTGCATCTGC 1045
 |||||
 QY 1449 ACAAGCAAGCCCGGAGAGAGTGTGAGAGGACACCTCTGTGCGAGACACATCG 1508
 |||||
 DB 1046 ACAAGCAAGCCCGGAGAGAGTGTGAGAGGACACCTCTGTGCGAGACACATCG 1105
 |||||


```
gene
CDS
    /clone_id="686 (synonym: hicc3). Vector pSPORT1_Sfi; host
    DH10B; sites SfiI + SfiIR"
    /dev_stage="adult"
    /note="fibronectin 1, differentially spliced"
    1..7777
    /gene="DKFZp686K08164"
    266..7069
```

```

1. 1.7777
2. "name="DKFZp686k08164"
3. 266. 7069
4. /gene="DKFZp686k08164"
5. /codon_start=-1
6. /product="hypothetical protein"
7. /protein_id="CAH18171.1"
8. /db_xref="GI:51476362"
9. /translation="MRGGSGPGLLLALVLCGLAVASTKSKSRPAEAYTPQSPVAV
10. VSGSRGCGCTDKNKHQIINOQMERTYIAGNALVCTCYGSGRPFNCSSKRAEAYTCDBKTYLT
11. ENTIRVGDYTERPKSMIDVCTCTAGRGRIISCTIANRCHGGQSYKIGDWRPHETG
12. GSGMILECVCLCKGEMWTKPKIAEKFCDHAAGSSYYVGEETQYQOMWMDCTCLSESG
13. GYATCTCSLRNKCNDQDTSTYKIGTMSKQNRGMLDICI CTNGNGEWKCEBCEG
14. OTTSSGSGPFTVRAVAYOPHPQRPYGHCTGSGVSVYGMQILKTQGNKMLCTC
15. CTNGNSGCEFTAVTQYGGNSNGERCVLPPTNGRPTVCTCTBGRDGHMLCSTTSN
16. EODKYSFCTDTHVLTQTRGNSNGMLCHPPLVNNHNYVDCTCEBGRDNKMKGGTYV
17. NYADQKQKFCFMAAHEICITNNEGAVYAGDMQDQMDGHMMKCTCVSGNGREBNT
18. IAYSDQKQICVDITVYNNADITPKHEBHEBMLCTCFQGRGKMKCDPYQDCSDST
19. GTFYDIGSMKEKAGSVKRWCKCYGSGIGMKHCOPLQTYSSSGSPVYLTETSPQNO
20. SHPLNOAMPOBHSIKYLLRNKPKSVGKKATCTHGLNSYTKLKKPVTYEGQLL
21. SIQYGHQGVTRPDTTSTSTPVSNTYVGETTPPSPLAVDTSEIYIATSEVSVSM
22. VSNASTVSGFRVREYIESEGDVQDIDPSTASVAPDILPKRIYVNYQISDEBES
23. QSLISTGCTGAPADAPEDPQVQVDSTIYVMSRQDAPITKRIYVSPVSGSTPES
24. MLPEANSTVSLDPOGYOYNTIYVAENQESTPVIQOETGTPESDPVSYPRIQD
25. FVEVDVYVLTIMTDEPVSIVAGRVVITPVNIPGHGQRLISRNTPAYTGLSPGTY
26. YTKFPAVSHGSRKELTRQCTTKDAPNLDQVNETDSTVLYRMTPRMOITGYRLT
27. VGLTRKQDPRQITVIVTWPAPRIQFKPVRPOCGEAPREYTSDBGSIYVSGLTPE
28. SIIPEYNEQEVNTIYVITWPAPRIQFKPVRPOCGEAPREYTSDBGSIYVSGLTPE
29. VEYATYIOLVLRGGNSLDAIVNKVAVPLSPPTMLLEANDTGLVMSRGETPDIIT
30. GRIITTPPGQOGBSLAEVHVAQDSCFEDNISPLGLENVSVYVYQDKESVTPIDPT
31. IIEVPQULTDLSFVDITDSIGLRKTPPLANSSTIIGRIYVAAAGGIIPIEDVYDSSV
32. GYIYVATGSEIGDIDYISVTLINGGSAAPTLTQVAVPEPDLRFTNLIGPDMRVSSV
33. AAPPSIDILTNELVRSVPMNEEDVAVLSIPDNVAVTLNLGTEVYVSVSYVEQH
34. ESTPLRGKQCTLDSPGTDFSDIYANFSEHILAPRATITGRIIRHHPGSRPRE
35. RVRHVSNSILTNLTPREYVSIYALNGRFSPLIQGQSYVDVPRILEVYAPAR
36. TSLLSMPAPAVTAYVYRTTYGETGNSPYQETVYGSKTATISGLKREVDYITIVY
37. AVTNGGSGPASKPISINRTIEDKSPQVYVDVQUNSIYVKMLPSSSPYGRKAVTNY
38. PRNGGPKPKTKAPQDQEMTEIGLOPQVYVSVVAONSQSGSPQVAVTAYI PAP
39. TDLKFTQVTPSLISAKMTPNQLTGYRIVRPEKGTGPKKEINLADSSSVYVGLM
40. VAKTEKSVYALIKDILTSAPGVVTLTNVSPRRARVATDAHETITISMRKETIT
41. TGPOVDAPANQGPQIQRITKIDVSSYITGLOPQYDYKIYLTATLNDMASSPVIDA
42. STALDAPSNRFLATVNSLWSQPPRARITVILIKTEPSSGPSPREVVPRPSTGEY
43. ATTGLBEGTETVYIALLKNQKSPPLGRKTKQGBALQSOTIISNAPFDQSTYIIS
44. CHAPGTDELPQFRPGVSTASTLITLGTAGTANILVEALKDQOHRHREVVYVYV
45. VNBGNIQPTDLSCPDPTYSYSHVAGDEMERSMSGKPLCCGCFSGSGHRRCOSSRNC
46. HDNVCATISCTCFGGQGRGRCNCRPAGSPPTPEGTGGSYNOISQKHORTINTVANC
47. KEYIACATISCTCFGGQGRGRCNCRPAGSPPTPEGTGGSYNOISQKHORTINTVANC
48. IIECFMPDLVQADRDSRR"

```

ORIGIN

Query Match	21.4%	Score 330;	DB 4;	Length 7777;
Best Local Similarity	100.0%	Pred. No. 2.5e-82;		
Matches 330; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1209	CCCATAGCTGGAAGTGTTTGATCAATGCTCTGGGATTCCTATATGATGTGGGAAACG	1268
Db	806	CCCATAGCTGGAAGTGTTTGATCAATGCTCTGGGATTCCTATATGATGTGGGAAACG	865
QY	1269	TGGGAGAGCCCTTACCAAGGCTGATGATGTGTACTTTCCTGGAGAGGAGCG	1328
Db	866	TGGGAGAGCCCTTACCAAGGCTGATGATGTGTACTTTCCTGGAGAGGAGCG	925
QY	1329	GGAGCGATCACTTGCATTTTAGAATATGATGCACGATCAGACCAAGGACATCTTAT	1388
Db	926	GGAGCGATCACTTGCATTTTAGAATATGATGCACGATCAGACCAAGGACATCTTAT	985
QY	1389	AGAAATTGAGACACCTGGAGCAAGAGATATATGAGAAACCTGCTCCATGCAATCTGC	1448
Db	986	AGAAATTGAGACACCTGGAGCAAGAGATATATGAGAAACCTGCTCCATGCAATCTGC	1045

QY	1449	ACAGGCAACGGCCGACGAGAGTGGAAAGTGTGACAGGACACCTCTGTGACAGACCAATCG	1508
Db	1046	ACAGGCAACGGCCGACGAGAGTGGAAAGTGTGACAGGACACCTCTGTGACAGACCAATCG	1105
QY	1509	AGCGATCTGGCCCTCTACCGAAGTTCGT	1538
Db	1106	AGCGATCTGGCCCTCTACCGAAGTTCGT	1135

[illegible]

REFERENCE
AUTHORS
CONSRM
TITLE
JOURNAL

1 (bases 1 to 7868)
Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,
Oanger, A., Fobo, G., Han, M. and Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (20-JAN-2005) MIPS, Ingolstaedter Landstr.1, D-85764

COMMENT
Clone from S. Wiemann, Molecular Genome Analysts, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; sequenced by GBR (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686M04163) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686M04163>
Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686M04163Q"
/db_xref="taxon:9606"
/clone="DKFZp686M04163"
/tissue_type="uterus endometrial, primary cell culture"
/clone_id="686 (synonym: hicc3). Vector pSPORT1_Sfi; host
DH10B; sites SfiI + SfiII"
/dev_stage="adult"
/note="fibronectin precursor"
1..7868

```

CDS

```

/product="hypothetical protein"
/protein_id="CAE45714.1"
/db_xref="GI:34364617"
/db_xref="GOA:Q6NOA6"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"
/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/trEMBL:Q6NOA6"
/translation="MRGPEPGLLAVLCTGAVPSTGASKSRQAOQMVQPSVA
VSSKSGCYDNKRYVINOQWERTYLGALVCTCYGSGSRFNCSPKEAEFTFDYT
GNTYRVGDIYERKDSIMWDCITIGAGRGISCTIANRCHGGQSYKIDYTPRPHET
GGYMLAEVCYLGNNGKEWTCCKPIAEKCFDAAGSYVVGWPEKPYGGMMVVDCTIGE
GSGRIITCRNRKNDODITRISYIGDTSKDKDNRGLLOCTITGNGRGSKCERHTSV
QTTSSSGGPTDRAVAVQPPHDPYPGHCVTDGSGVYSVGMWLKTYGNKNKMLCT

```


Qy	1209	CCCATAGCTGAGAGTGTTTTGATCATGCTGCTGGGACATTCCTATGTGGTTCGAGAAACG	1268
Db	808	CCCATAGCTGAGAGTGTTTTGATCATGCTGCTGGGACATTCCTATGTGGTTCGAGAAACG	867
Qy	1269	TGGGAGAAGCCCTACCAAGCGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAAGCGACG	1328
Db	868	TGGGAGAAGCCCTACCAAGCGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAAGCGACG	927
Qy	1329	GGACGCATCACTTGCATCTTCTAGAAAATAGATAGCAACGATCAGGACACAAGGACATCCTAT	1388
Db	928	GGACGCATCACTTGCATCTTCTAGAAAATAGATAGCAACGATCAGGACACAAGGACATCCTAT	987
Qy	1389	AGAATTGGAGACACTTGGAGCACAAGAGGATAATCGAGAAAACCTGCTCCAGTGCATCTGC	1448
Db	988	AGAATTGGAGACACTTGGAGCACAAGAGGATAATCGAGAAAACCTGCTCCAGTGCATCTGC	1047
Qy	1449	ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG	1508
Db	1048	ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG	1107
Qy	1509	AGCGGATCTGGCCCCCTTACCGATGTTTCGT	1538
Db	1108	AGCGGATCTGGCCCCCTTACCGATGTTTCGT	1137

RESULT 14
AI095589/c
LOCUS
DEFINITION
AI095589 465 bp mRNA linear EST 05-OCT-1998
gb24408.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
IMAGE:1697174 3' similar to gb:X02761_cds1 FIBRONECTIN PRECURSOR
(HUMAN); mRNA sequence.
ACCESSION
VERSION AI095589.1 GI:3434565
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
TITLES 1. (bases 1 to 465)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
INRAT Consortium (info@image.llnl.gov) for further information.
Insert Length: 1871 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 446.
Location/Qualifiers
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1697174"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/note="Organ: uterus; Vector: p7T73-Pac; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
AATCGGAGATTCGGCGCGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."

Query Match 21.3%; Score 329; DB 1; Length 465;

	Best Local Similarity	99.7%; Pred. No. 2.2e-82;							
	Matches 329; Conservative	0; Mismatches	1; Indels	0; Gaps	0; Gaps	0; Gaps	0; Gaps	0; Gaps	0; Gaps
Qy	1209	CCCATAGCTCAGAAGTGTGTTTTGATCATGCTCGTCGGGACATTCCTATGTGTGGTCGAGAAACG	1268						
Db	434	CCCATAGCTCAGAAGTGTGTTTTGATCATGCTGCTGGGACTTCCTATGTGTGGGAGAAACG	375						
Qy	1269	TGGGAGAAGCCCTACCAAGGCTGATGATCGTAGATTGTACTTTGCCCTGGGAGAAGGCAGC	1328						
Db	374	TGGGAGAAGCCCTACCNAAGCTGNATGATGGTAGATTGTACTTTGCCCTGGGAGAAGGCAGC	315						
Qy	1329	GGACGCATCACTTTGCACTTCTAGAAAATAGATGCACAACGATCAGGACACAAAGGACATCCTAT	1388						
Db	314	GGAAGCATCACTTTGCACTTCTAGAAAATAGATGCACAACGATCAGGACACAAAGGACATCCTAT	255						
Qy	1389	AGAAITGGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAAGTGCACTGC	1448						
Db	254	AGAAITGGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAAGTGCACTGC	195						
Qy	1449	ACAGGCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTTGTGTGCAGACCAACATCG	1508						
Db	194	ACAGGCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTTGTGTGCAGACCAACATCG	135						
Qy	1509	AGCGGATCTGGGCCCTTACCGATGTTTGGT	1538						
Db	134	AGCGGATCTGGGCCCTTACCGATGTTTGGT	105						

RESULT	15
AU140802	
LOCUS	734 bp mRNA linear EST 05-AUG-2002
DEFINITION	AU140802 PLACE4 Homo sapiens cDNA clone PLACE4000266 5', mRNA sequence.
ACCESSION	AU140802
VERSION	AU140802.1 GI:11002323
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 734) Ota.T., Suzuki.Y., Saito.K., Iehii.S., Yamamoto.J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuko.Y. and Isogai,T. HRI human cDNA project (Ota.T., Suzuki.Y., Saito,K., Iehii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuko,Y., Isogai,T.) Unpublished (2000) Contact: Takao Isogai
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

Location/Qualifiers
1. .734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000266"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pMR1ASFL3"

FEATURES
source

ORIGIN

Query Match 21.3%; Score 329; DB 1; Length 734;
Best Local Similarity 99.7%; Pred. No. 2.5e-82;
Matches 329; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```

QY 1209 CCTATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGACTTCTATGTGTGGAGAAACG 1268
    |||||
Db 309 CCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGACTTCTATGTGTGGAGAAACG 368
    |||||
QY 1269 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGTGTATGATTGTACTTGCTGGAGAAAGCAGC 1328
    |||||
Db 369 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGTGTATGATTGTACTTGCTGGAGAAAGCAGC 428
    |||||
QY 1329 GGACGCATCATTTGCACTTCTAGAAATAGATGCAAGATCAGAGACACAGGACATCTCTAT 1388
    |||||
Db 429 GGACGCATCATTTGCACTTCTAGAAATAGATGCAAGATCAGAGACACAGGACATCTCTAT 488
    |||||
QY 1389 AGAATTGAGAGACACTGAGACAAAGAGATTAATCGAGAAACCTGCTCCAGTGCACTGCG 1448
    |||||
Db 489 AGAATTGAGAGACACTGAGACAAAGAGATTAATCGAGAAACCTGCTCCAGTGCACTGCG 548
    |||||
QY 1449 ACAAGCAATGCGCCGAGAGAGATGGAAGTGTGAGAGGACACCTCTGTGCAGACCACATCG 1508
    |||||
Db 549 ACAAGCAATGCGCCGAGAGAGATGGAAGTGTGAGAGGACACCTCTGTGCAGACCACATCG 608
    |||||
QY 1509 AGCGGATCTGGGCCCTTCACCGATGTTCTG 1538
    |||||
Db 609 AGCGGATCTGGGCCCTTCACCGATGTTCTG 638
    |||||

```

Search completed: February 1, 2006, 12:22:47
 Job time : 6336.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:47:34 ; Search time 107.195 Seconds
(without alignments)
12632.698 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 2716
Sequence: 1 ttgttctaacttaagaag.....ccttcacgatgtcgttag 1541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 segs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2.1/USFTO_spool_p/US09940235/runat_27012006_144217_27548/app.query.fasta_1.7708
-DB=A.GeneSeq -OFMT=faстан -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1_1_846 @runat_27012006_144217_27548 -NCPU=6 -ICPU=3
-NO MPAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
A_GeneSeq_21:*
1: geneSeqp19808:*
2: geneSeqp19908:*
3: geneSeqp20008:*
4: geneSeqp20018:*
5: geneSeqp20028:*
6: geneSeqp20038:*
7: geneSeqp200308:*
8: geneSeqp20048:*
9: geneSeqp20058:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984	73.0	413	8	ADL92189
2	1984	73.0	414	2	AAR10194
3	1984	73.0	414	2	AAR63120
4	1984	73.0	414	2	AAW24794
5	1984	73.0	414	2	AAW94664
6	1984	73.0	414	3	AAV01556
7	1984	73.0	795	2	AAV4797
8	1984	73.0	795	2	AAV4797
9	1981	72.9	531	2	AAR11829

10	1980	72.9	414	2	AAW94665	AAW94665 Streptococ
11	1978	72.8	483	2	AAR12522	AAR12522 Factor Xa
12	1966.5	72.4	813	2	AAW21728	AAW21728 Wild type
13	1966.5	72.4	1194	2	AAW21726	AAW21726 Streptoki
14	1963	72.3	415	2	AAW93425	AAW93425 Streptoco
15	1958	72.1	414	2	AAW86143	AAW86143 Streptoki
16	1958	72.1	414	3	AAW01295	AAW01295 Wild type
17	1958	72.1	483	3	AAR12889	AAR12889 Streptoki
18	1958	72.1	483	2	AAR12885	AAR12885 Factor Xa
19	1958	72.1	499	2	AAR12891	AAR12891 Streptoki
20	1958	72.1	859	2	AAR12893	AAR12893 OmpA
21	1954	71.9	414	2	AAR20202	AAR20202 S.equisim
22	1951.5	71.9	415	3	AAV50870	AAV50870 Streptoco
23	1951.5	71.9	415	3	AAV99593	AAV99593 Streptoco
24	1947	71.7	414	2	AAW86144	AAW86144 De-immun
25	1947	71.7	414	3	AAW01296	AAW01296 Altered s
26	1942	71.5	440	2	AAW06377	AAW06377 Streptoki
27	1937.5	71.3	1194	2	AAW21725	AAW21725 Modified
28	1936.5	71.3	1194	2	AAW21724	AAW21724 Modified
29	1928	71.0	413	2	AAV25020	AAV25020 Streptoki
30	1928	71.0	413	5	ABB80012	ABB80012 Streptoki
31	1928	71.0	413	6	ABG74199	ABG74199 Represent
32	1924	70.8	372	2	AAR10200	AAR10200 Streptoki
33	1916	70.5	372	2	AAR10197	AAR10197 Streptoki
34	1907.5	70.2	371	2	AAR10195	AAR10195 Streptoki
35	1907	70.2	374	2	AAR10198	AAR10198 Streptoki
36	1904	70.1	414	4	AAW74940	AAW74940 Mutant st
37	1902	70.0	384	3	AAW84007	AAW84007 Amino aci
38	1900	70.0	800	2	AAW21723	AAW21723 Plasmidog
39	1900	70.0	1181	2	AAW21727	AAW21727 Streptoki
40	1898	69.9	372	3	AAW84006	AAW84006 Amino aci
41	1890.5	69.6	747	2	AAR12894	AAR12894 Met-core
42	1887	69.5	401	3	AAW84004	AAW84004 Amino aci
43	1887	69.5	413	3	AAW84005	AAW84005 Amino aci
44	1875	69.0	369	2	AAR12892	AAR12892 Truncated
45	1855	68.3	440	1	AAW50620	AAW50620 Sequence

ALIGNMENTS

RESULT 1
ADL92189
ADL92189 standard; protein; 413 AA.

AC ADL92189;
AC 20-MAY-2004 (first entry)
DT 20-MAY-2004 (first entry)
XX Streptokinase protein sequence.
DE Streptokinase protein sequence.
XX harvesting; recombinant; host cell; N-terminal leader peptide;
KW pre-peptide; lantibiotic; post-translational modification;
KW pharmaceuticals; vaccine; immunogenic.
XX
XX OS Unidentified.
XX WO2003099862-A1.
XX 04-DEC-2003.
XX PD 26-MAY-2003; 2003WO-NL000389.
XX PF 24-MAY-2002; 2002EP-00077060.
XX PR 07-FEB-2003; 2003US-00360101.
XX (NANO-) APPLIED NANOSYSTEMS BV.
XX Mol1 GN, Leenhouts CJ, Kuipers OP, Driessen AJM;
XX WPI; 2004-042770/04.
XX Harvesting a desired polypeptide produced by a recombinant host cell, for
PT producing pharmaceuticals, comprises selecting a recombinant nucleic acid

comprising nucleic acid fragments encoding a leader peptide and the polypeptide.

Claim 4; Page 82-83; 109pp; English.

The invention relates to a novel method for harvesting a (poly)peptide produced by a recombinant host cell. The novel method involves selecting a cell comprising a first nucleic acid encoding a leader peptide and a second nucleic acid fragment encoding the desired (poly)peptide. The first and second fragments are within the same open reading frame of the first nucleic acid and the leader peptide is functionally equivalent to an N-terminal leader peptide found with the pre-peptide of a lantibiotic. The host cells and nucleic acids are useful for producing, harvesting and post-translational modification of polypeptides. The polypeptides may be used in the production of pharmaceuticals, e.g. as antigen for vaccine or immunogenic composition. This sequence represents a polypeptide relating to the novel method of the invention.

Sequence 413 AA;

Alignment Scores:

Pred. No.: 1.13e-178 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 8 Gaps: 0

US-09-940-235-9 (1-1541) x ADL92189 (1-413)

QY	33	ATAGTCGTCTCGAATGGCTACTAGATCGTCTCTCTGTAATAACACCAATTTGGTGT	92
DB	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
QY	93	AGCGTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC	152
DB	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	40
QY	153	GATCTAAATCATCGACCTGCTCATGGAGGAAAGACAGACAGGCTTAAGTCCAAATCA	212
DB	41	AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer	60
QY	213	AAACCATTTGCTACTGATGTCGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA	272
DB	61	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu	80
QY	273	AAGGCTATTCAAGAACAAFTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC	332
DB	81	LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	100
QY	333	ATTGATTTTGCAGGATGCAACCAATTACTGATCGAAACGGCAGGCTACTTTGCTGAC	392
DB	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
QY	393	AAAGATGTCGTCGTAACCTTCGCGACCCCACTGTCCAAAGAATTTTGTAAAGCGACAT	452
DB	121	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	140
QY	453	GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG	512
DB	141	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	513	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTGACAGGCTCTCAA	572
DB	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
QY	573	GATACCTAGCTATTGAAACACTAGCTATCGGTGACCATCACATCTCAAGAACTACTA	632
DB	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu	200
QY	633	GCTCAAGCACAAGCATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACCTGAC	692
DB	201	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	220

QY	693	TCCTCAATCGTCACTCATGCAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG	752
DB	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	753	TTTACTTACCGTGTAAAAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	812
DB	241	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	260
QY	813	AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	872
DB	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
QY	873	GAAGAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATACGTT	932
DB	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
QY	933	GATGTCGATACCAACGAATTTGTAAGAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC	992
DB	301	AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
QY	993	TTAGACTTCAGAGATTTATACGATCTCTCGTGAATAAGCTAAAGCTACTCTACAAATCTC	1052
DB	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	340
QY	1053	GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGCTAGAGGATATACGATGAC	1112
DB	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1113	ACCAACGCTATCATACCCGTTTATATGGCAAGCCAGCCCAAGAGAGAGAATGCTAGCTAC	1172
DB	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
QY	1173	CATTAGCT 1181	
DB	381	HisLeuAla 383	
RESULT 2			
ID	AA10194	standard; protein; 414 AA.	
XX	AA10194;		
AC	AA10194;		
DT	28-MAR-1991	(first entry)	
XX			
DE		Streptokinase encoded by synthetic gene.	
XX		streptokinase; thrombolytic agent; myocardial infarction.	
KW			
XX		Synthetic.	
PN	EP407942-A.		
PD	16-JAN-1991.		
XX			
PF	11-JUL-1989;	89JP-00179432.	
XX			
PR	11-JUL-1989;	89JP-00179432.	
XX			
PR	27-NOV-1989;	89JP-00307957.	
XX			
PR	11-APR-1990;	90JP-00096830.	
XX			
PA	(SAKA) OTSUKA PHARM FACTOR.		
XX			
PI	Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;		
XX			
DR	WPI; 1991-016179/03.		
XX	N-PSDB; AAQ10230.		
DR			
XX			
PT		Synthetic gene encoding streptokinase - scale, high purity prodn. of	
XX		streptokinase used as a thrombolytic agent.	
PS	Claim 1; Page 51; 76pp; English.		
XX			

CC Streptokinase and its derivatives can be produced in large quantities
CC with high purity for use as thrombolytic agents in patients with lung
CC thrombus or myocardial infarction. See also AAR10195-R10200

XX Sequence 414 AA;

Alignment Scores:

Pred. No.:	Length:	414
Score:	1.13e-178	Matches: 383
Percent Similarity:	1984.00	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	73.05%	Indels: 0
	2	Gaps: 0

US-09-940-235-9 (1-1541) x AAR10194 (1-414)

```

QY 33 AATGCTGGCTCTGAAGGCTACTAGATCGCTCTTCTGTAATAACAGCCAAATGGTTGT 92
Db 1 IleaIagLyProGIuIrPleuLeuAspArgProSeValAsnAsnSerGIuLeuVal 20
QY 93 AGCGTTGCTGTAAGCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 152
Db 21 SerValaIagLyThrValaGIuGIuIYrThrAsnGlnAspIleSerLeuYsPhePheGIuIle 40
QY 153 GATTAACATCAAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAAGTCCAAATCA 212
Db 41 AspLeuThrSerArgProAlaHisGIyGIyLysThrGIuGIuIYleuSerProIYsSer 60
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATPAACTTGAGAAAGCTGACTTA 272
Db 61 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGIuYsAlaAspLeuLeu 80
QY 273 AAGGCTATCAAGAACAAATTGATCGTAACTGCTACAGATCAAGACCTAATTGAGTGC 332
Db 81 LysAlaIleGIuGIuIYleuIleAlaAsnValHisSerAsnAspArgIYrPheGIuVal 100
QY 333 ATTTGATTTTGAAGCGATCAACCATTAAGTCAAGACGAGCTCTTGGCTGAC 392
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGIyYsValIYrPheAlaAsp 120
QY 393 AAAGATGGTTCGGTAACTTGGCGAGCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 452
Db 121 LysAspGIySerValThrLeuProThrGIuProValGIuGIuIYleuLeuSerGIyHis 140
QY 453 GTGCCCGGTGAACCATATAAGAAAAACCAATPACAAACCAAGCAATTTGTTGATGTG 512
Db 141 ValArgValaIArgProIYrLysGIuYsProIleGIuAsnGlnAlaYsSerValaAspVal 160
QY 513 GAATATACGTGTACAGTTTACTCCCTTAAACCTGATGAGATTTTGACAGCGGTCAAA 572
Db 161 GIuIYrThrValaGIuPheThrProLeuAsnProAspAspPheArgProGIyLeuYs 180
QY 573 GATACCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGATTACTA 632
Db 181 AspThrIYsLeuLeuYsThrIleuAlaIleGIyAspThrIleThrSerGIuIYleuLeu 200
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCAAGGCTATACGATTATGAACGTGAC 692
Db 201 AlaGIuAlaGIuIYleuSerIleuAsnYsAsnHisProGIyYrThrIleYrGIuArgAsp 220
QY 693 TCCCTCAATCGTCAATCAATGACAAATGCAATTTTCCGTAACGATTTTACCATGATCAAGAG 752
Db 221 SerSerIleValaIYrHisAspAsnAspIlePheArgThrIleLeuProMetAspGIuGIu 240
QY 753 TTTACTTACCGGTGTAAATAATCGGGAAAGGCTTATAGATCAATPAAAAATCGTGTG 812
Db 241 PheThrIYrArgValaIYsAsnAsnArgGIuGIuAlaIYrArgIleAsnYsYsSerGIyLeu 260
QY 813 AATGAAGAAATTAACAAACATGACCTGATCTCTGAGAAATATTACGTCCTTAATAAAAGG 872
Db 261 AsnGIuGIuIleAsnAsnThrAspLeuIleSerGIuIYsYrYrValleuYsYsGIy 280
QY 873 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACATCAATACGTT 932

```

```

Db 281 GIuYsProIYrAspProPheAspArgSerHisLeuYsLeuPheThrIleYsYrVal 300
QY 933 GATGTCATATCCAAAGCAATTCCTTAAAGTAGAGCTCTTAACAGTCAAGCAATGAC 992
Db 301 AspValaAspThrAsnGIuLeuLeuYsSerGIuGIuIYleuThrAlaSerGIuArgAsn 320
QY 993 TTAAGCTTCAGAGATTATATGATCCCTGATAGGCTTAAGGCTTAAGTCAACAAATCTC 1052
Db 321 LeuAspPheArgAspLeuYrAspProArgAspYsAlaYsLeuLeuYrAsnAsnLeu 340
QY 1053 GATGCTTTGGTATATATGACTATACCTTAACTGAAAGAGTAGAGATATACAGATGAC 1112
Db 341 AspAlaPheGIyIleMetAspIYrThrLeuThrGIySValaGIuAspAsnHisAspAsp 360
QY 1113 ACCAAGCGTATCATTAACCGTTTATATGGCAAGCGCAAGCAAGAGATATGCTACTAC 1172
Db 361 ThrAsnArgIleIleThrValaIYrMetGIyLysArgProGIuGIuAsnAlaSerYr 380
QY 1173 CATTAGCT 1181
Db 381 HisLeuAla 383

RESULT 3
AAR63120
ID AAR63120 standard; protein; 414 AA.
XX
AC AAR63120;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 16-NOV-1994 (first entry)
XX
DE Streptokinase.
XX
KM Streptokinase; SK; Streptococcus equisimilis; plasminogen;
KW myocardial infarction.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
FH Key
FT Region
FT /note= "claim 3, see CC"
FT Region
FT /note= "claim 1, see CC"
FT Region
FT /note= "claim 3, see CC"
FT Region
FT /note= "claim 3, see CC"
FT Region
FT /note= "claim 3, see CC"
FT Region
FT /note= "claim 2, see CC"
PN W09407992-A1.
XX
PD 14-APR-1994.
XX
XX
XX 05-OCT-1993; 93WO-US009502.
XX
XX 05-OCT-1992; 92US-00956692.
XX
XX 29-SEP-1993; 93US-00128299.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX (HARD ) HARVARD COLLEGE.
XX
XX
XX Reed GL;
XX
XX WPI, 1994-135561/16.
XX
XX DNA encoding a polypeptide which binds to plasminogen and corresponds to
XX PT region of streptokinase - useful to detect plasminogen in a sample and to
XX PT treat myocardial infarction.
XX
XX Disclosure; Page 40-41; 62pp; English.
PS

```

XX Nucleic acid comprising a sequence encoding amino acids 14-414 of
CC streptokinase, which binds to plasminogen and does not have a sequence
CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 414 AA;
Alignment Scores:
Pred. No.: 1.13e-178 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x AAR63120 (1-414)
QY 33 ATAGCTGGTCCGCTGCTAGTCTCTCTGTAATAACAGCCAAATGGTGTGTT 92
DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 93 AGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 152
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
QY 153 GATCTAACATCAGCAGCTGCTCATGGAGGAAGACAGACAGGCTTAAGTCCAAATCA 212
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAAACGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 272
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeu 80
QY 273 AAGGCTATTCAAGAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 332
DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 333 ATTGATTTTGCAGGATGCAACCATTAATGATGCAAGACGCGAAGTCTACTTGTG 392
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAAGATGTTCCGTTACCTTCGCGCCCAACCTGCTCAAGAAATTTTGTCTAAGCG 452
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTG 512
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATCTGTACGTTACTCCCTTAAACCTGTATGACGATTCAGACAGGTCTCAA 572
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATATACGCTATTCAACACATGCTATCGGTGACACCATCACATCTCAAGNATTACTA 632
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
QY 633 GCTCAAGCACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAC 692
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATGCTCCTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTACTTACCGTGTAAAAATCGGACGACGCTTATAGATCATATAAAAAATCTGCTCG 812
DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 813 AATGAAGAAATAAACACACCTGACCTGATCTCTGAGAAATATTTACCTCTCTTAAAAAGGG 872

DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCCGCTATGATCCCTTTGATCGCAGTCACCTGAAACCTGTTCAACATCAATAGCTT 932
DB 281 GluLysProTyrAspPropheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGCTGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 992
DB 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGCTGATAAGGCTAACTCTCAACAATCTC 1052
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTGCTGATTTGCTGATCTTAACTGGAAGCTAGAGGATATACGATGAC 1112
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACCGTATACACCGTTTATATGGCAAGCGACCCGAGAGAGAGATGCTAGCTAC 1172
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
DB 381 HisLeuAla 383
RESULT 4
AA24794
ID AAY24794 standard; protein; 414 AA.
XX
AC AAY24794;
XX
DT 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
DE Streptococcus equisimilis native streptokinase.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nSK; rSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN WO9931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US026694.
XX
PR 15-DEC-1997; 97US-0069497P.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1999-395183/33.
DR N-PSDB; AAX80492.
XX
PT N-terminally deleted streptokinase.
XX
PS Claim 30; Page 60-61; 73pp; English.
XX
CC The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase; (3) an
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a

CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents native streptokinase (nsk). (Updated on 17-OCT-2003 to
 CC standardise OS field)

XX Sequence 414 AA;

Alignment Scores:

Pred. No.:	1,13e-178	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x AAY24794 (1-414)

QY 33 ATAGCTGCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCAATTGGTTGTT 92
 Db 1 IleaIaGIyProGIuITrleuLeuAaPaRgProSeValaAsnHsSerGIuLeuVal 20
 QY 93 AGCGTGTCTGTACTGTGAGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 152
 Db 21 SerValaIaGIyThrValaGIuGIyThraSngIaAspIleSerleuIySphPheGIuIle 40
 QY 153 GATCTTAACATCAACACCTCTCATGAGAGAAACAGAGAGAGCTTAATCCAAATCA 212
 Db 41 AspleuIThrSerArProAlaHsGIyGIySthGIuGIuGIyLeuSerProIySer 60
 QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAGCTGACTTA 272
 Db 61 LysPProPheAlaIThrAspSerGIyAlaMetSerHsIySleuGIuIySalaAspleuLeu 80
 QY 273 AAGGCTATTCAAGAACAATTGATCGCTAAAGTCCACAGTACAGCACTTCTTGAGTTC 332
 Db 81 LysAlaIaIeGIuGIuGIuLeuIleAlaAsnValHsSerHsAspIyRrPheGIuVal 100
 QY 333 ATTGATTTTGGCAGGATGCAACCAATTACTGATGGAAGCGCAAGCTTACTTGTCTGAC 392
 Db 101 IleAspPheAlaSerAspAlaIThrIleIThrAspArgAsnGIyValITyrPheAlaAsp 120
 QY 393 AAAGATGGTTCGGTAACTTGGCGAGCCCAACCTGTCACAGAAATTTTGTAAACGGACAT 452
 Db 121 LysAspGIySerValIThrleuProIThrGlnProValGIuGIuIySleuLeuSerGIyHs 140
 QY 453 GTGCGCGTTAGACCATATTAAGAAAAACCAATCAAAACCAAGCAAACTGTGTGATGTC 512
 Db 141 ValArgValArgProITyRlySgluIySproIleGIuHsGIuHsAlaIySserValaAspVal 160
 QY 513 GAATTAAGTGAAGATTATCTCCCTTAAACCTGATGACGATTTCAGACCGAGTCTCAAA 572
 Db 161 GluITyThrValaGIuHsPheIThrProleuAsnProAspAspPheArgProGIyLeuIyS 180
 QY 573 GATATTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTCACTTCAAGAAATTA 632
 Db 181 AspITrIySleuLeuIySthIleuAlaIeGIyAspITrIleIThrSerGIuGIuLeuLeu 200
 QY 633 GCTCAAGACAAAGCATTTTAAACAAACCAACCCGAGGCTATTCGATTTATGAACGTGAC 692
 Db 201 AlaGIuHsAlaGIuSerIleuLeuSnyAsnHsAspProGIyITyRrIleITyGIuIyAsp 220
 QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTCAATGATGATCAAG 752
 Db 221 SerSerIleValIThrHsAspAsnAspIlePheArgIThrIleuIThrProMetAspGIu 240

QY 753 TTATCTTACCGTGTAAATAATCGGAAACACCTTATAGATCAATAAAAATCGTCTG 812
 Db 241 PheITrITyRArgValIySAsnHsArgGIuIuIaITyRArgIleAsnIySleuSerGIyLeu 260
 QY 813 AATGAAGAAATTAACACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAG 872
 Db 261 AsnGIuGIuIleAsnAsnIThrAspIleuIleSerGIuIySITyRITyValleuIySgl 280
 QY 873 GAAAGCCGTATGATCCCTTGTATCGGAGTCACTTGAACCTTGAACCAATCAATACCTT 932
 Db 281 GluIySProITyRAspProPheAspArgSerHsIySleuIySleuPheITrIleITyVal 300
 QY 933 GATGTCATACCAACGAATTGCTTAAAAAGTACGAGCTCTTAACAGTAGCAACGTAAAC 992
 Db 301 AspValaAspIThrAsnGIuLeuLeuIySserIuGIuLeuIThrAlaSerGIuArgHs 320
 QY 993 TTGACCTTCAGAGATTATTAACGATCTCTGTATTAAGCTTAACTTACCAACATCTC 1052
 Db 321 LeuAspPheArgAspLeuITyRAspProArgAspIySalaIySleuLeuITyRAsnHsLeu 340
 QY 1053 GATGCTTTGGTATTAAGACTATACCTTAACCTGAAAAGTAGAGATAATACAGATGAC 1112
 Db 341 AspAlaPheGIyIleMetAspITyRIThrleuIThrGIySValaGIuAspAsnHsAsp 360
 QY 1113 ACCAACGGTATCATACCGTTTATATGAGCAAGCAACCGAAGAGAGATGCTAGCTAC 1172
 Db 361 ThrAsnArgIleIThrITrValITyMetGIyAspArgProGIuGIuIyAsnHsAspITyR 380
 QY 1173 CATTTAGCT 1181
 Db 381 HisIleuAla 383

RESULT 5
 AAW94664
 ID AAW94664 standard; protein; 414 AA.
 AC AAW94664;
 XX
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis native streptokinase.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note="encoded by ACC"
 PN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR N-PSDB; AAX16632.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.

PS Claim 1; Col 7-10; 17pp; English.
 XX The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence represents native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPLg) to
 CC plasmin (HPLm), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX

SQ Sequence 414 AA;

Alignment Scores:
 Pred. No.: 1.13e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x AAW94664 (1-414)

QY	33	ATAGCTGCTCTGATGGCTACTAGATCGTCTCTTCTGTAATAACAGCCAAATGGTGT	92
Db	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
QY	93	AGCGTTGCTGTACTGTGGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC	152
Db	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	40
QY	153	GATCTAACATCACGACCTCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA	212
Db	41	AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	60
QY	213	AAACATTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA	272
Db	61	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu	80
QY	273	AAGGCTATTCAAGAACATTGATCGCTACGTCACAGTAAACGACGACTACTTTGAGTTC	332
Db	81	LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	100
QY	333	ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC	392
Db	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
QY	393	AAAGATGTTTGGTAACTTCGTCGACCCCACTCTGTCACAGAAATTTTGTAAACGACAT	452
Db	121	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	140
QY	453	GTGCGGTAGACCATATAAGAAACCAATACAAACCAACGCGAATCTGTGATGTCG	512
Db	141	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	513	GAATATATCTGACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGGCTCTCAA	572
Db	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys	180
QY	573	GATACTAGCTATTGAAACACTAGTATCGGTGACACCATCACATCTCAAGAAATCTA	632
Db	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	200
QY	633	GCTCAAGCACAAAGCATTTAAACAAACCAACCCAGGCTATACGTTTATGACGTCAC	692
Db	201	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	220
QY	693	TCCTCAATCGTCACTCATGACAAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG	752

Db	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	753	TTTACTTTACCGTGTAAAAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	241	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	260
QY	813	AATGAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG	872
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
QY	873	GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAAAATAGCTT	932
Db	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
QY	933	GATGTCGATACCAACGAATTTCTAAAAGTGACGAGCTCTTAACAGCTAGCGACGTAAC	992
Db	301	AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
QY	993	TTAGACTTCAGAGATTTTATACGATCCTCGTGATAAGGCTAAACTTACTCTACAACAATCTC	1052
Db	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	340
QY	1053	GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCACGATGAC	1112
Db	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1113	ACCAACGATATCATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGCTAC	1172
Db	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
QY	1173	CATTTAGCT 1181	
Db	381	HisLeuAla 383	

RESULT 6
 AAY01556
 ID AAY01556 standard; peptide; 414 AA.
 XX
 AC AAY01556;
 XX
 DT 17-OCT-2003 (revised)
 DT 18-JUN-1999 (first entry)
 XX
 DE Native streptokinase protein sequence.
 XX
 KW Antigenic peptide; streptokinase; streptokinase-specific antibody;
 KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
 KW rheumatic fever.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 PN WO9908698-A1.
 XX
 PD 25-FEB-1999.
 XX
 PF 18-AUG-1998; 98WO-US017114.
 XX
 PR 18-AUG-1997; 97US-0055911P.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Reed GL, Parhami-Seren B;
 DR WPI; 1999-190113/16.
 XX
 PT New polypeptides which bind streptokinase-specific antibodies - useful in
 PT thrombolytic therapy.
 XX
 PS Disclosure; Page 12; 44pp; English.
 XX
 CC The present sequence represents a native streptokinase. The specification
 CC describes a polypeptide which binds to a streptokinase-specific antibody

CC and prevents the antibody binding to native streptokinase. The
 CC specification also describes a synthetic polypeptide (PI) comprising an
 CC epitope which binds to an streptokinase-specific antibody and reduces
 CC thrombolytic activity of streptokinase. PI is used in thrombolytic
 CC therapy and to prevent or treat glomerulonephritis and rheumatic fever.
 CC (Updated on 17-Oct-2003 to standardise OS field)

XX Sequence 414 AA:

Alignment Scores:

Pred. No.:	1 13e-178	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x AA01556 (1-414)

```

OY 33 ATAGCTGCTCGAATGGCTACTAGATCGTCTTGTAAATAACAGCAATGGTGT 92
DB 1 11ealaglyProdlutlrpleuleaapargProSerValashnenserGlnleuVal 20
OY 93 AGCGTTGCTGTACTGTGAGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 152
DB 21 SerValalaglyThrValGluGlyThrAsnGlnaspIleSerleuyspnehegIule 40
OY 153 GATTTAAACATCAGACCTGCTCATGAGGAAAGACAGCAAGCTTAAGTCCAAATCA 212
DB 41 AspLeuThrSerArgProalashlglyGlyThrGlnGlnGlyLeuSerProlySer 60
OY 213 AAACATTTGCTACTAGTAGTGGCGGATGTCATTAACCTTGAGAAAGCTTACTA 272
DB 61 LysProPhealathraspserGlyalameSerHisLysleuGluValaAspLeu 80
OY 273 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAACGACGACTCTTGAAGTC 332
DB 81 LysAlaIleGlnGlnleuIleAlaenValHisSerashnasparyrPheGluVal 100
OY 333 ATTTGATTTTGCAGGACGACACCATTTACTGATCGAAACGGCAAGCTTACTTGTGAC 392
DB 101 IleaspPhealaseraspalathrIleThrAspArgasnGlyValTyPheAlaAsp 120
OY 393 AAAGATGTTGGTAACTTGGCCGACCACTGTCACAGAAATTTTGTAAACGGACAT 452
DB 121 LysAspGlySerValThrleuProthGlnProValGlnGlnleuLeuSerGlyHis 140
OY 453 GTGCGGTTAGACCATATTAAGAAAAACAATACAAACCAAGCGAAATCTGTGATGTG 512
DB 141 ValArgValArgProIlyrlysgIuysProIleGlnAsnGlnAlaLysSerValaAspVal 160
OY 513 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTGACAGCGGTCTCAA 572
DB 161 GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
OY 573 GATCTAAGCTATTTGAAAAACACTAGCTATGGTGACACCATCTCAAGAAATTTCTA 632
DB 181 AspThrLysleuLeuysrThrleuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
OY 633 GCTCAAGCAACAAGCTTTTAAACAACCAACCAAGCGGCTATCGATTTATGACAGGAC 692
DB 201 AlaGlnAlaGlnSerIleleuasnLysasnhsPProGlyTyThrIleTyrgluAsp 220
OY 693 TCCCTCAATGTCATCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAG 752
DB 221 SerSerIleValThrHisaspasnaspIlePheArgThrIleleuProketaAspGlnGlu 240
OY 753 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTATAGATCATATAAAAAATCTGGTCTG 812
DB 241 PheThrTyArgValLysasnArgGlnAlaTyArgIleasnLysLysSerGlyLeu 260
OY 813 AATGAAGAAATAACAACCTGACCTGATCTCGAAGAAATATATACGCTTAAAAAAGG 872

```

```

DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluysTyTyValleuLysGly 280
OY 873 GAAAGCGGTATNGATCCCTTTGATCGGACGTACATGAAACGTTACCATTAACGTT 932
DB 281 GluysProIlyrAspProPheaspArgSerHisleuysleuPheThrIleTyVal 300
OY 933 GATGTCGATCAACAAGAAATGCTAAAGAGAGCAGCTTTACAGCTAGCAACGTAC 992
DB 301 AspValaspThrAsnGlnleuLeuLysSerGlnleuLeuThrAlaSerGluArgAsn 320
OY 993 TTAGACTTCAGAGATTTATACGATCTCGTATTAAGGCTAAACCTTACACAAATCTC 1052
DB 321 LeuaspPheArgAspLeuTyAspProArgAspLysAlaLysleuLeuTyAsnAsnleu 340
OY 1053 GATGCTTTGGTATTTATGAGACTTATACCTTAACGTGAAAAAGTATACAGATGAC 1112
DB 341 AspAlaPheGlyIleMetAspTyThrleuThrGlyValGluAspAsnHisaspAsp 360
OY 1113 ACCAAGCTATCATACCGTTTATATGAGGCAAGGACCGCAAGAGAGAGATGCTAC 1172
DB 361 ThrasnArgIleIleThrValTyMetGlyLysArgProGluGlyluAsnAlaSerTy 380
OY 1173 CATTTAGCT 1181
DB 381 HisleuAla 383

```

RESULT 7
 AA090282
 ID AA090282 standard; protein; 414 AA.
 XX
 AC AA090282;
 DT 12-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 S. equisimilis streptokinase.
 DE
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 EN EPI024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RES.
 PI Sahní G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 DR WPI; 2000-516032/47.
 DR N-PSDB; AAA37633.
 XX
 PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 PS Example 3; Fig 3; 58pp; English.
 XX
 CC This sequence represents the human Streptococcus equisimilis
 CC streptokinase protein sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprising a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (PBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a

CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 414 AA;

Alignment Scores:
 Pred. No.: 1.13e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0
 DB: 3 Gaps: 0

US-09-940-235-9 (1-1541) x AAY90282 (1-414)

QY	33	ATAGCTGCTCTGAATGGCTACTAGATCGCTCTCTCTGTAATAACAGCCAAATGGTGT	92
DB	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	20
QY	93	AGCTTGTCTGCTACTGTTGAGGGGACGATCAAGACATTAGTCTTAATTTTGAATC	152
DB	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	40
QY	153	GATCTAATCACCAGCTCTCATGAGGAGAAAGACAGACAGGCTTAAGTCCAAATCA	212
DB	41	AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer	60
QY	213	AAACCATTTGCTACTGATGAGCGCGATGTCACATAAATTTGAGAAAGCTCACTACTA	272
DB	61	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu	80
QY	273	AAGCTATTTCAGACAAATTCCTGCTAAGCTCCAGTCAAGACGACTACTTGGTGC	332
DB	81	LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	100
QY	333	ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGCAAGTCTACTTTGTGAC	392
DB	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
QY	393	AAAGATGTTCCGTTAACTTCGCGACCCCAACCTGTCCAGAAATTTTGTAAAGCGACAT	452
DB	121	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	140
QY	453	GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGG	512
DB	141	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	513	GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACGAGTCTCAA	572
DB	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
QY	573	GATACTAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA	632
DB	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	200
QY	633	GCTCAGACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTGAC	692
DB	201	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	220
QY	693	TCCTCAATCGTCACTCAGCAATCACATTTTCCGTACGATTTTACCATGATCAGAG	752

DB	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	753	TTTACTACCGTGTAAATAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG	812
DB	241	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	260
QY	813	AATGAAGAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	872
DB	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly	280
QY	873	GAAGACCTATGATCCCTTTGATCGAGTCACCTTGAACCTGTTCACCATCAATACGTT	932
DB	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
QY	933	GATGTCGATACCAACCAATTTGCTTAAAGTGAAGTACGAGCTCTTAACAGCTACGACGTAAC	992
DB	301	AspValAspThrAsnGluLeuLysSerGluGlnLeuThrAlaSerGluArgAsn	320
QY	993	TTAGACTTCAGAGATTTATACGATCTCTCGTGAAGCTAACTACTCTACAACAATCTC	1052
DB	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu	340
QY	1053	GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAGGATATACGATGAC	1112
DB	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1113	ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGCTAC	1172
DB	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
QY	1173	CATTTAGCT 1181	
DB	381	HisLeuAla 383	

RESULT 8
 AAY24797
 ID AAY24797 standard; protein; 795 AA.
 XX
 AC AAY24797;
 XX
 DT 26-AUG-1999 (first entry)
 XX
 DE Streptokinase and maltose binding protein fusion protein.
 XX
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW rSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 PN W09931247-A1.
 XX
 PD 24-JUN-1999.
 XX
 PF 15-DEC-1998; 98WO-US026694.
 XX
 PR 15-DEC-1997; 97US-0069497P.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Reed GL;
 XX
 DR WPI; 1999-395183/33.
 XX
 DR N-PSDB; AAX80497.
 XX
 PT N-terminally deleted streptokinase.
 XX
 PS Example; Page 48-51; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical

CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents a streptokinase and maltose binding protein fusion protein
 CC from an example of the present invention

XX Sequence 795, AA:

Alignment Scores:
 Pred. No.: 1,48e-178 Length: 795
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x AAY24797 (1-795)

QY 33 ATAGCTGGTCCCTGAATGCTACTGATCGTCTCTGTAATTAACAGCCAAATGGTTGTT 92
 Db 382 IleaIaGlyProGIuItrPleuLeuAerProSeValaAsnAserGlnIleuVal 401
 QY 93 AGCGTTGCGTACTGCTGAGGGGGAAGATCAAGACATTAGCTTAATTTTGAATC 152
 Db 402 SerValaIaGlyThrValaGlnGlyItrAenGlnAerPIeSerLeuYsPhepegIuIle 421
 QY 153 GATTTACATCAGACCTGCTCATGAGGAAAGACAGACGCTTAAGTCCAAATCA 212
 Db 422 AspleuThrSerAtrProIahIseGlyIyIyThGlnGlnIlyLeuSerProIySer 441
 QY 213 AAACCATTTGCTACTGATAGTGCGCGGATGTCACATTAACCTTGAGAAAGCTGACTTA 272
 Db 442 LysProPheAlaThrAerSerGlyAlaMetSerHleIyLeuGluIySalaAerLeuLeu 461
 QY 273 AAGGCTATTCAAGAACAAATTAATCGCTAACGTCACAGTAACAGACCTACTTGAAGTC 332
 Db 462 LysAlaIleGlnGlnGlnIleuIleAlaenValHleSerAynAerAerProIyPheGluVal 481
 QY 333 ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAAGCGCAAGGCTACTTGGCTAC 392
 Db 482 IleAerPheAlaSerAerPalathrIleThrAerAerAerGlyIySValIyTrPheAlaAer 501
 QY 393 AAAAGATGGTTCGGTAACTTGGCGAACCAACCTGTCCAAGAAATTTTGTACCGGACAT 452
 Db 502 LysAerPGLySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHls 531
 QY 453 GTGGCGGTTAGACCATATAAAGAAAAACAAATCAAAACCAAGCGAAATCTGTGATGTG 512
 Db 522 ValArgValaIarProIyTrIySglIySProIleGlnAenGlnAlaIySerValaIarVal 541
 QY 513 GAATATACGTACAGTTTATCTCCCTTAACCCGATGATGAGATTTCGACACAGGCTCGAA 572
 Db 542 GlnYTrItrValaGlnPheThrProLeuAenProAerAerAerPheAerProGlyLeuLeu 581
 QY 573 GATATAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCATCAAGATTATCA 632
 Db 562 AspThrIySleuLeuLeuYerThrIleuAlaIleGlyAerThrIleThrSerGlnIleuLeu 581

QY 633 GCTCAGACCAACCAATTTTAAACAAAAACCAACCGGCTATAGCAATTTATGAACGTAC 692
 Db 582 AlaGlnIaGlnSerIleLeuAerLysAerHlsProGlyTrItrIleIyGlnAerAer 601
 QY 693 TCTTGAAATCGTACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 752
 Db 602 SerSerIleValThrHlsAerAerAerPIePheAerGThrIleLeuProMetAerGlnIu 621
 QY 753 TTTACTTAACCGTGTAAATAAACCGGGAACACCTTAATGGAACATTAATAAATTCGTCTG 812
 Db 622 PheThrIyTrArgValIySaerAerGlnGlnAlaIyTrArgIleAerLysLysSerGlyLeu 641
 QY 813 AATGAAGAAATTAACCAACACTGACCTGATCTGAGAAATATTAAGTCTTAATAAAGGG 872
 Db 642 AsnGlnIuIleAenAenThrAerLeuIleSerGlnIySlyTrIyValIleuIySGLy 661
 QY 873 GAAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTACCATCAATACGTT 932
 Db 662 GlnYsProIyTrAerProPheAerAerSerHlsLeuYsLeuPheThrIleIySlyVal 681
 QY 933 GATGTCATACCAACGAATTCCTTAATAAAGTACAGACGCTTAAACGCTAGCAACGTAC 992
 Db 682 AspValaAspThrAenGlnLeuLeuLysSerGlnIleuLeuThrAlaSerGlnAerAer 701
 QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGTAAGCTTAACCTCAACAAATCTC 1052
 Db 702 LeuAerPheAerAerLeuYTrAerProIyAerAerAerAerAerAerAerAerAer 721
 QY 1053 GATGCTTTGGTATTAATGACCTATACCTTAACCTGAAAGTGAAGATTAATCAAGTAC 1112
 Db 722 AspAlaPheGlyIleMetAerAerIyThrLeuThrIlySValGlnAerAerHlsAerAer 741
 QY 1113 ACCAAGCGTATCATTAACCGTTTATATGGGCAAGGACCCGAAGAGATGCTAGTAC 1172
 Db 742 ThrAerAerGlyIleIleThrValIyMetGlyLysAerProGlnGlnIyGlnAerAer 761
 QY 1173 CATTTAGCT 1181
 Db 762 HleLeuAla 764

RESULT 9
 ID AAR11829
 XX AAR11829 standard; protein; 531 AA.
 AC AAR11829;
 XX 08-JUL-1991 (first entry)
 DT
 DE FB-FB-SK fusion conjugate.
 XX Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
 KW streptokinase; fusion protein.
 XX
 OS Staphylococcus aureus.
 XX
 XX Key Location/Qualifiers
 FH Peptide 2..61
 FT Peptide /label= FB monomer
 FT Peptide 62..119
 FT Peptide /label= FB monomer
 FT Peptide 120..531
 FT /label= streptokinase
 PN
 XX
 XX US5011686-A.
 PD 30-APR-1991.
 XX
 XX 15-NOV-1989; 89US-00437769.
 PF
 XX 21-SEP-1987; 87US-00099242.
 PR
 XX (CREA-) CREATIVE BIOMOLEC.
 PA
 XX

PI Pang RHL;
 XX WPI; 1991-140198/19.
 DR N-PSDB; AAQ11651.
 XX
 XX Imparting injectable fibrinolytic agent - with affinity for intravascular
 PT thrombus, by linking agent to fibrin binding domain.
 XX
 XX Disclosure; Fig 5; 18pp; English.
 PS
 XX
 CC The conjugate comprises an FB-FB dimer linked to streptokinase The FB
 CC fragment has selective affinity for fibrin, low affinity for fibrinogen,
 CC and minimal immunogenicity, imparting thrombus- targeting capability.
 CC See also AAR11821 and AAR11828
 XX
 XX
 SQ Sequence 531 AA;
 Alignment Scores:
 Pred. No.: 2,41e-178 Length: 531
 Score: 1981.00 Matches: 382
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 72.94% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-9 (1-1541) x AAR11829 (1-531)
 QY 33 ATAGCTGCTCCTGAATGGCTACTAGATCGTCTCTCTGTAATAACAGCCAAATGGTGT 92
 DB 118 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnThrGlnLeuVal 137
 QY 93 AGCGTTGCTGGTACTGTTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTTGAAATC 152
 DB 138 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 157
 QY 153 GATCTAACATCAGCACTGCTCATCGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
 DB 158 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 177
 QY 213 AAACATTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA 272
 DB 178 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 197
 QY 273 AAGGCTATTCAAGAACATTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC 332
 DB 198 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 217
 QY 333 ATTGATTTTGCAGCGATGCAACCACTTACTGATCGAAACGCGAAGGCTCTACTTTGCTGAC 392
 DB 218 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 237
 QY 393 AAAGATGGTTGGTAACTTGGCGGACCCAACTGTCCTCAAGAAATTTTGTAGCGGACAT 452
 DB 238 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 257
 QY 453 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATG 512
 DB 258 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 277
 QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTCAGACGAGTCTCAAA 572
 DB 278 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 297
 QY 573 GATACATAGCTATTGAAACACTAGCTATCGGTGACCAATCACAATCTCAAGAAATCTA 632
 DB 298 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 317
 QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTCAC 692
 DB 318 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 337
 QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 752

DB 338 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 357
 QY 753 TTTACTTACCGTGTATAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGCTG 812
 DB 358 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysSerGlyLeu 377
 QY 813 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
 DB 378 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 397
 QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACACTGTTACCATCAATAGCTT 932
 DB 398 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 417
 QY 933 GATGTCGATACCAACGAATTTGCTAAAAGTGACAGCTCTTAACAGCTACGACGTAAC 992
 DB 418 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 437
 QY 993 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAAGCTTAAACTTACTCTCAACAATCTC 1052
 DB 438 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 457
 QY 1053 GATGCTTTGCTATTATGACTATACCTTAACCTGGAAGTAGAGGATAATCAGATGAC 1112
 DB 458 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 477
 QY 1113 ACCAACCGTATCATACCGCTTTATATGGAAGCGACCGACCGAGAGAGATGCTAGCTAC 1172
 DB 478 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 497
 QY 1173 CATTTAGCT 1181
 DB 498 HisLeuAla 500
 RESULT 10
 AAW94665
 ID AAW94665 standard; protein; 414 AA.
 XX
 AC AAW94665;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis mutant streptokinase K59E.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 14 /note= "encoded by ACC"
 FT
 FT
 XX US5876999-A.
 XX
 XX 02-MAR-1999.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 XX (NASC-) NAT SCI COUNCIL.
 XX
 XX Wu H;
 PI
 XX WPI; 1999-189643/16.
 DR N-PSDB; AAX16633.
 XX
 XX Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT

DR WPI; 1991-208151/28.
 XX N-PSDB; AAQ12490.
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 XX prophylaxis.
 PS Disclosure; Page 96; 115pp; English.
 XX
 CC The protein is a recombinant product of a gene fusion construct. The
 CC sequence of the synthetic hirudin HV-1 genes was designed based on the
 CC published amino acid sequence (Dodd J., et al. FEBS Letters 165 180
 CC (1984)). The sequence of streptokinase was obt'd. from PCR amplified
 CC chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
 CC used for the PCR were based on the pub- lished DNA sequence of S.
 CC equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
 CC 357-362 [1985]). The two sequences were used to construct an expression
 CC vector in which the streptokinase gene is linked to the hirudin gene via
 CC a linking sequence encoding a cleavage site for factor Xa. The factor Xa
 CC is present at the site of the target thrombus so the active agents are
 CC released specifically at the place where clot formation is occurring. See
 CC also AAR12887-R12889, AAR12891-R12894 and AAR12885. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX
 SQ Sequence 483 AA;

Alignment Scores:
 Pred. No.: 4,46e-178 Length: 483
 Score: 1978.00 Matches: 401
 Percent Similarity: 82.28% Conservative: 17
 Best Local Similarity: 78.94% Mismatches: 52
 Query Match: 72.81% Indels: 38
 DB: 2 Gaps: 8

US-09-940-235-9 (1-1541) x AAR12522 (1-483)

QY 33 ATAGTGTGCTGAATGGCTACTAGATCGTCTCTGTAATAAACAAGCAATTTGGTTGT 92
 DB 1 IleAlaGlyProGluTrpLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 153 GATCTAACATCAGCAGCTGCTCATGGAGGAAGACAGACAGCTTAAGTCCAAATCA 212
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 213 AAACCATTTGCTAGTAGTGCGCGCATGTCCATATAAATCTTGAGAAAGCTGACTTACTA 272
 DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
 QY 273 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAACCGACTACTTTGAGGTC 332
 DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 333 ATTGATTTGCAAGCGATGCAACCAATTACTATGATCGAAACGCGAAGTCTACTTTGCTGAC 392
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 393 AAAGATGGTTCGGTAACCTTGCAGACCAACCTGTCCAGAAATTTTGTAAAGCGACAT 452
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 453 GTGCGGCTTAGACCATATAAGAAAAACCAATACAAAAACAAGCGAAATCTGTTGATGTG 512
 DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 513 GAATATACGTACAGTTTACTCCCTTAAACCTTGATGAGATTTTCAGACAGGTCTCAA 572
 DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 QY 573 GATACTAGCTATTGAAACACACTACTATCGGTGACACCATCATCTCAAGNAATTACTA 632

DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 633 GCTCAAGCACAAAGCATTTTAAACAAAAACCAAGCAGCTATACGATTATGACGTCAC 692
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgasp 220
 QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTCAGATTTTACCAATCGATCAAGAG 752
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 753 TTTACTTACCGTGTAAAAATCGGAACAAGCTTTATAGGATCAATAAAAAATCTGGCTG 812
 DB 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
 QY 813 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 872
 DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCAACATCAATACGTT 932
 DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 992
 DB 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 993 TTGACTTTCAGAGATTTATACGATCTCTCGTGAATAGGCTAACTCTCTACAAATCTC 1052
 DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1053 GATGCTTTGCTGATATGAGTATACCTTAACCTGGAAGCTAGAGGATTAATCAGATGAC 1112
 DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1113 ACCAACCGTATACAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAC 1172
 DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1173 CATTTAGCTGGTGGCCAGCGCCAAACAGATTGTACCCATGCTGAGAAAGTGTGTTGAT 1232
 DB 381 HisLeuAla-----TyrAspLysAspArgTyrThrGluGluAlaArgGluValTyrSer 398
 QY 1233 CATGCTGCTGGGACTTCTATGTCGAGAAACGTCGGGAGAGAGCCCTACCAAGGCTGG 1292
 DB 399 TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLysIleGluGlyArg 418
 QY 1293 ATGATG---GTAGATTGCTACT-----TGCCCTGGGAGAGAGCGAGCGGA 1331
 DB 419 ValValTyrThrAspCysThrGluSerGlyGlnAsnLeuCysLeuCysGluGlySerAsn 438
 QY 1332 CGCATCACTTGCATCTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTATAGA 1391
 DB 439 -----ValCysGlyGlnGlyAsnLysCys-----TGCCCTGGGAGAGAGCGAGCGGA 1447
 QY 1392 ATTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCACA 1451
 DB 448 LeuGlySerAspGlyGluLysAsn-----GlnCysVal---Thr 459
 QY 1452 GGCACACGCGGAGGAGAGTGGAGGCGACACCTCTGTGTCAGACACCATCTCGAGC 1511
 DB 460 GlyGluGly-----ThrProLysProGlnSerHis 469
 QY 1512 GGATCTGGCCCTTTCACCGATGTT 1535
 DB 470 AsnAspGlyAspPheGluGluIle 477
 RESULT 12
 AAW21728
 ID AAW21728 standard; protein; 813 AA.
 XX
 AC AAW21728;
 XX

DT		17-OCT-2003 (revised)
DT		01-OCT-1997 (first entry)
XX		
XX		Wild type plasminogen-binding fragment of Streptokinase.
DE		
XX		Plasminogen-binding fragment; streptokinase; degradation; MBP;
KW		thrombolytic agent; blood clot; bolus; maltose-binding protein.
KW		
XX		Streptococcus dysgalactiae subsp. equisimilis.
OS		
XX		WO9641883-A1.
PN		
XX		27-DEC-1996.
PD		
XX		
PE		07-JUN-1996; 96WO-US009640.
PR		09-JUN-1995; 95US-00488940.
XX		
PA		(HARD) HARVARD COLLEGE.
XX		
PI		Reed GL;
DR		WPI, 1997-065469/06.
XX		
PT		Modified forms of streptokinase resistant to enzymatic cleavage - useful
XX		as thrombolytic agents in treating thrombosis and in medical equipment.
XX		
PS		Example 1; Page 12-13; 65DP; English.
CC-		This sequence represents the wild type plasminogen-binding fragment of
CC		streptokinase. This fragment was used in the design of a modified
CC		streptokinase has an in vitro degradation rate at least 2 times slower
CC		than that of native streptokinase. Compounds containing modified
CC		streptokinases are specifically used as thrombolytic agents for
CC		dissolving blood clots in vivo in a mammal, preferably at a dose of 200mg/kg,
CC		opt. as a bolus rather than by continuous infusion. (updated on 17-
CC		OCT-2003 to standardise OS field)
SO		Sequence 813 AA;
XX		
Alignment Scores:		
Pred. No.:	6.83e-177	Length: 813
Score:	1966.50	Matches: 382
Percent Similarity:	99.74%	Conservative: 0
Best Local Similarity:	99.74%	Mismatches: 0
Query Match:	72.40%	Indels: 1
DB:	2	Gaps: 1
US-09-940-235-9 (J-1541) x AAM21728 (1-813)		
DY	33 ATAGTGTGCTCGAATGGCTACTAGATCGTCCTTGTAATAACAGCCAAATTGGTTGT	92
Db	1 IleaIaGIyProGIuNtRpeuleuAdPaRGProSeValaAmnSergInleuVal	20
DY	93 AGCGTTGCGTGACTGTTAGGGGAGCAATCAAGACAATTAGTCTTAATTTTTGAATC	155
Db	21 SerValaIaGIyThVaIGluGIyTranSngInAspIIeSerXeuysPhePhGIuile	40
DY	153 GATCAAACATCAAGCACTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAATGCCAAATCA	212
Db	41 AsrPeuInmSerArgProAlahIsGIyGLyStRhIngInImLyLeuSerProLysSer	60
DY	213 AAACCAATTTGCTACTGATAGTGCGGCAGATGTCACATAACTTGAGAAGCTGACTACTA	272
Db	61 LysPProPheIadlThAsrSergIylametSerHislyLeuGIuIyValaAsrLeuLeu	80
DY	273 AAGGCTATTCAAGACAATTTGATGCTGAACGTGCAAGTACAGACACTATTGAGGTC	332
Db	81 LysaIaIleGIingInImleuIeaIaaEnValHisSerXenAsrAspTyRPhaIaVal	100
DY	333 ATTGATTTTTGAAGGAGTGAACCACTTACTGATGCAAAAGCGCAAGGCTACTTTGCTGAC	392
Db	101 IleAsrPheIaISerAsrPaIatThrIIetmAsrPaGAndIyUyValTyRPhaIaAsp	120

QY	333	AAAGATGGTTGGGTAACCTTGCACGACCCAACTGTCCAAAGATTCTTAAGGGACAT	452
Db	121	LysAspGlySerValThrLeuProthGlnProValGlnGluPheLeuLeuSerGlyHis	140
QY	453	GTGCGCGTTGAAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	512
Db	141	ValArgValArg--TyrGlyGluSerProIleGlnAsnGlnAlaIleSerValAspVal	159
QY	513	GAATATATATGATACGTTTACTCCCTTAACCCCTGATGACATTTACAGCCAGGTCTCAA	572
Db	160	GluYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	179
QY	573	GATACATAAGCTAATGAAAAACCTAGCTATCCGTGTGACACCATCATCTCAAGATTACTA	632
Db	180	AspThrIryLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu	199
QY	633	GCTCAGACCAAGACGATTTTAAACAAAACCAACCCAGGCTATACGATTATGAACGTGAC	692
Db	200	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	219
QY	693	TCCCTCAATCGTCACCTGACAAATGACATTTTCCGTACGATTTTACCAATGACCAAG	752
Db	220	SerSerIleValAlaThrAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	239
QY	753	TTTACTTATCCGTGTAAAAAATCGGGAACAACTTATAGATCATATAAAATCTGCTGTG	812
Db	240	PheThrIryArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysSerGlyLeu	259
QY	813	AATGAAAGAAATAAACAACACTGACCTGATCTCTGAAATATTAACGTCTTAAAAAAGG	872
Db	260	AsnGluGluIleAsnAsnThrAspLeuIleSerGluSerTyrTyrValLeuLysGly	279
QY	873	GAAAAAGCGATATGATCCCTTTGATCGACGTCACTTGAACCTGTCACATCAATAAGCTT	932
Db	280	GluLysProIryAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	299
QY	933	GATGTGCATACCAACGAATGTCTAAAAAGTGACAGGCTTTAACAGCTAGCAACGTAAAC	992
Db	300	AspValAspThrAsnArgIleLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	319
QY	993	TTAAGACTTCAGAGATTTATACGATCTCTGTGATTAAGGCTAAACTACTTCACACATCTC	1052
Db	320	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAspAsnLeu	339
QY	1053	GATCCTTTGGTATTATGACTATACCTTAACGTGAAAAAGTAGAGATTAATCACGATGAC	1112
Db	340	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	359
QY	1113	ACCAACCGTATCATACACGTTTATATGAGGCAAGCAACCCGAAGAGAGAAATGCTAGCTAC	1172
Db	360	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr	379
QY	1173	CATTAGCT 1181	
Db	380	HisLeuAla 382	
RESULT 13			
ID	AAW21726	standard; protein; 1194 AA.	
XX	AAW21726;		
AC	17-OCT-2003	(revised)	
XX	01-OCT-1997	(first entry)	
XX	Streptokinase/maltose binding protein fusion protein, rSK.		
DE			
XX	Plasminogen-binding fragment; streptokinase; degradation; MBP;		
KW	thrombolytic agent; blood clot; bolus; maltose-binding protein.		
XX			
XX	Streptococcus dysgalactiae subsp. equisimilis.		

Key Location/Qualifiers
 1.381
 /label= Maltose binding protein.
 /note= "acts as 'blocking group'"
 382..1194
 /label= Streptokinase

W09641883-A1.

27-DEC-1996.

07-JUN-1996; 96WO-US009640.

09-JUN-1995; 95US-00488940.

(HARD) HARVARD COLLEGE.

Reed GL;

WPI; 1997-065469/06.

Modified forms of streptokinase resistant to enzymatic cleavage - useful as thrombolytic agents in treating thrombosis and in medical equipment.

Example 1; Page 11-12; 65pp; English.

This sequence represents a fusion protein between maltose-binding protein and the plasminogen-binding fragment of streptokinase. This fusion protein was used in the design of modified streptokinase has an in vitro degradation rate at least 2 times slower than that of native streptokinase. Compounds containing modified streptokinases are specifically used as thrombolytic agents for dissolving blood clots in vivo in a mammal, preferably at a dose of 2000 U/kg, optionally as a bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to standardise OS field)

Sequence 1194 AA;

Alignment Scores:

Pred. No.: 8.02e-177 Length: 1194
 Score: 1966.50 Matches: 382
 Percent Similarity: 99.74% Conservative: 0
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 72.40% Indels: 1
 DB: 2 Gaps: 1

US-09-940-235-9 (1-1541) x AAW21726 (1-1194)

QY	33	ATAGTCTGCTCGTAATGGCTACTAGATCGTCTCTGTAAATAACAGCCAAATGGTGT	92
DB	382	IleAlaGlyProGluTrpLeuAspArgProSerValAsnAsnSerGlnLeuValVal	401
QY	93	ACGCTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC	152
DB	402	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	421
QY	153	GATCAATACATCAGCCTGCTCATCGAGGAAGACAGCAAGCGTCTTAAGTCAAAATCA	212
DB	422	AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	441
QY	213	AAACATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA	272
DB	442	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu	461
QY	273	AAGGCTATTCAAGAACAAATTGATCGCTACGTCACAGTAACGACGACTACTTTGAGGTC	332
DB	462	LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	481
QY	333	ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC	392
DB	482	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	501
QY	393	AAAGATGGTTCGGTAACCTTGCCGACCCCAACCTGTCCAGAAATTTTGTGTAACGGACAT	452

DB	502	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	521
QY	453	GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTGTATGTG	512
DB	522	ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	540
QY	513	GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAA	572
DB	541	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	560
QY	573	GATATAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAATACTA	632
DB	561	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	580
QY	633	GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTGAC	692
DB	581	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	600
QY	693	TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	752
DB	601	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	620
QY	753	TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGTCTG	812
DB	621	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	640
QY	813	AATGAAGAATAAACAACACTGACCTGATCTCGAGAAATATTACGTCTTAAAAAGGG	872
DB	641	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	660
QY	873	GAAGCCGCTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCAACCAATACGTT	932
DB	661	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	680
QY	933	GATGTCGATACCAACGAATTTGCTAAAAAGTGACAGCTCTTAAACAGCTACGCAACGTAA	992
DB	681	AspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	700
QY	993	TTAGACTTCAGAGATTTTATACGATCTCTCGTGAATAAGCTTAACTCTCTCAACAATCTC	1052
DB	701	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	720
QY	1053	GATGCTTTGCTATTATGCTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC	1112
DB	721	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	740
QY	1113	ACCAACCGTATCATACCCGTTTATATGCGCAACGACCCCGAAGGAGAGATGCTAGCTAC	1172
DB	741	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	760
QY	1173	CATTAGCT 1181	
DB	761	HisLeuAla 763	

RESULT 14

AEB93425
 ID AEB93425 standard; protein; 415 AA.

XX
 AC AEB93425;

XX
 DT 20-OCT-2005 (first entry)

DE Streptococcus equisimilis streptokinase protein SEQ ID NO:1.

XX
 KW protein quantitation; streptokinase.

XX
 OS Streptococcus dysgalactiae subsp. equisimilis.

XX
 PN US2005176085-A1.

PD 11-AUG-2005.

XX

PF 15-NOV-2004; 2004US-00988943.
XX
PR 19-NOV-2003; 2003CU-00000269.
XX (BETA/) BETANCOURT N L H.
XX (DORT/) DORTA-DUQUE J F D C.
PA (PERE/) PEREZ V A B.
PA (VALD/) VALDES J G.
PA (LOPE/) LOPEZ L J G.
PA (PALO/) PALOMARES G R P.
PA (FEYT/) FEYT R P.
PA (GILF/) GIL F M A.
XX
PI Belancourt NLH, Dorta-Duque JFD, Perez VAB, Valdes JG, Lopez LJG;
PI Palomares GRP, Feyt RP, Gil FMA;
XX
DR MPI; 2005-596348/61.
XX
PT Identifying and quantifying proteins in complex mixtures by selectively
PT isolating peptides not containing histidine nor arginine from each
PT protein, and determining the relative concentration of proteins in
PT different samples.
PS
PS Example 3; SEQ ID NO 1; 91pp; English.
XX
XX The invention relates to a method for identifying and quantifying one or
CC more proteins in complex mixtures by selectively isolating peptides not
CC containing histidine nor arginine (NHR peptides) from each protein, and
CC determining the relative concentration of one or more proteins in
CC different samples from the ratio between the areas of estimated
CC theoretical spectra for each NHR peptide labeled with different isotopes
CC in each sample. The method comprises: (a) enzymatically or chemically
CC hydrolyzing the sample or samples of proteins; (b) chemically modifying
CC alpha and epsilon amino groups (alpha- and epsilon-NH2) of every peptide
CC obtained in step (a); (c) isolating the NHR peptides by cation exchange
CC chromatography from the mixture of peptides obtained in step (b); (d)
CC identifying proteins by mass spectrometry analysis of the NHR peptides
CC obtained in step (c); (e) differential isotopic labeling of protein
CC samples previously to step (a) or during steps (a) or (b) and immediately
CC mixing at least a portion of the samples; and (f) relative quantifying of
CC more proteins in the mixtures of step (e) from the ratio between the
CC areas of estimated theoretical spectra of the pair of NHR peptides
CC identified in step (d), as well as from the ratio between the areas of
CC the estimated theoretical spectra of fragments from the NHR peptides,
CC generated in step (d). Also described is a kit for the identification and
CC quantification of proteins in complex mixtures, which comprises the
CC method mentioned above. The method and kit are useful for identifying or
CC quantifying proteins, such as those with vaccinal, therapeutic or
CC diagnostic uses, in complex mixtures. The present sequence represents a
CC streptokinase protein, which is used in an example from the present
CC invention.
XX
XX Sequence 415 AA;
SQ
Alignment Scores:
Pred. No.: 1,116-176 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 72.28% Indels: 0
DB: 9 Gaps: 0
US-09-940-235-9 (1-1541) x AEB93425 (1-415)
QY 30 ATGATAGTGTGCTCTGAATGGCTACGATCGCTTCTGTAAATAACAGCAATTGGT 89
Db 1 MetilealaglyProglutirpleuleuaspargProSerValasnhsenrglnleVal 20
QY 90 GTTACGTTGCTGTGTACTGTGAGGAGCAATCAAGACTTAAGTCTTAATTTTGA 149
Db 21 ValSerValalaglyThrValgluglyThraenGlnaspIleSerleuylsPhepGlu 40
QY 150 ATCGATCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAA 209

Db 41 IleaspleuthrSerArgProAlahisgllygluylsThrGlnGlnGlyLeuSerProlys 60
QY 210 TCAGAAACATTTGCTACATGATAGTGGCGGATGTCACATTAACCTTGAAGAAAGCTTAA 269
Db 61 SerlyspProPhealathrAspsercllyAlaMetProhlslylsleuGlnlyalaSpleu 80
QY 270 CTAAGGCTATTCAGAAACAATTAAGCTTAAGCTCAAGCTCAAGTCAAGTCAAGCTTGA 329
Db 81 LeuylsalaileglnGlnGlnleuilelaasnValhisSerAsnAspPtyrPheGlu 100
QY 330 GTCAATGATTTGCAAGGATGCAACCATTAATGATGCAAAACGGCAAGTCTACTTGTCT 389
Db 101 ValilleaspPhealaSerAspAlathrIlethrAspArganglyllyValtyrPheala 120
QY 390 GACAAAGATGTTGGGTACCTTGGCCAGCCCAACTGTTCAGAAATTTTGGTAAAGCGA 449
Db 121 AspLysaspGlySerValThrleuProthrGlnProValGlnGlnPheleuSerGly 140
QY 450 CATGGGCGCTTAGCCATATTAAGAAACCAATACAAACCAAGGAAATCTGTGAT 509
Db 141 HisValargValargProtyrLysGlnLysProIleGlnaenGlnAlaLysSerValasp 160
QY 510 GTGGAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGAGTCTC 569
Db 161 ValglutyrThrValGlnPheThrProleuasnProAspAspAspPheargProgluLeu 180
QY 570 AAAATATCTAAGCTATTGAAAAACACTGCTATCGGTGACACCATCAATCAAGATTA 629
Db 181 LysaspThrlylsleuLeuylsThrleualaileglysPthrIleThrSerGlnGluLeu 200
QY 630 CTACTCAAGACAAAGCATTTTAAACAAACCAAGGCTTACGATTTAAGAAAGCT 689
Db 201 LeuulaglnalaglnSerIleleuasnlyshrhlsProgluLtyrThrIletyrGlnLys 220
QY 690 GACTCCTCAATCGTCACTCATGACATGACATGACATTTTCCGTAGCAATTTTACATGATCAA 749
Db 221 AsperSerIleValthrHisaspAsnAspIlePheargThrIleleuProMetAspGln 240
QY 750 GAGTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGCATCAATAAAAATCTGT 809
Db 241 GluPheThrlyrHisVallylsAsnArgGlnGlnAlaLtyrGlnIleasnlylsSerclly 260
QY 810 CTGAATGAAGAAATTAACACATGACCTGATCTCTGAGAAATTTAGTCTCTTAAAAA 869
Db 261 LeuasnGlnGlnleuasnAsnThrAspLeuIleSerGluylsTyrTyValleuLyslys 280
QY 870 GGGGAAAGCGGTATGATCCCTTGTATCGAGTCACTTGAACCTTACCATCAAAATAC 929
Db 281 GlyGlnLysProtyrAspProPheAspArgSerHisleuLysleuPheThrIleLysTy 300
QY 930 GTTGATGTCATACCAAGCAATTGCTTAAAAAGTAGAGAGCTCTTAACAGCTAGCGAAGT 989
Db 301 ValaspValasnThrAsnGlnleuLeuLysSercllyGlnleuLeuThrAlaSerGlnArg 320
QY 990 AACTTAAGCTTCAGAGATTTATACGATCTCTGATTAAGGCTAAATCTTACACAAAT 1049
Db 321 AsnleuAspPheargAspLeuTyAspProArgAspLysAlaLysleuLeuTyAsnAsn 340
QY 1050 CTCGATCTTTGGTATTATTAAGCATTAACCTTAACGAAAGAGTAGAGATTAACAGAT 1109
Db 341 LeuaspAlaPheGlyIleMetAspTyThrleuThrGlnLysValGlnAspAsnHisasp 360
QY 1110 GACACCAACCGTATCATTAACCGTTTATATGAGGCAAGGCAACCGAAGAGAGAAATGCTAGC 1169
Db 361 AspThrAsnArgIleIleThrValtyrMetcllyAspArgProGlnGlnAsnHisSer 380
QY 1170 TACCATTTAGCT 1181
Db 381 TyrtHisleuala 384
RESULT 15
AAM86143

AAW86143 standard; protein; 414 AA.
AAW86143;
17-OCT-2003 (revised)
03-MAR-1999 (first entry)
Streptokinase (SK) protein sequence.
Non-immunogenic; epitope; T-cell; immunogenicity; immune system; SK;
immunogl bulin; therapeutic; streptokinase.
Streptococcus dysgalactiae subsp. equisimilis.
WO9852976-A1.
26-NOV-1998.
21-MAY-1998; 98WO-GB001473.
21-MAY-1997; 97GB-00010480.
31-JUL-1997; 97GB-00016197.
28-NOV-1997; 97GB-00025270.
02-DEC-1997; 97US-0067235P.
14-APR-1998; 98GB-00007751.
(BIOV-) BIOVATION LTD.
Carr FU;
WPI; 1999-045301/04.
Reducing immunogenicity of proteins - by modifying the amino acid
sequence of the protein to eliminate potential epitopes for T-cells of a
given species.
Example 6; Fig 28; 77pp; English.
The invention relates to a method for the production of non-immunogenic
proteins. The method comprises determining at least part of the amino
acid sequence of the protein; (b) identifying in the amino acid sequence
one or more potential epitopes for T-cells (T-cell epitopes) of the given
species; and (c) modifying the amino acid sequence to eliminate at least
one of the T-cell epitopes identified in step (b) thereby to eliminate or
reduce the immunogenicity of the protein when exposed to the immune
system of the given species. A method of analysing a pre-existing protein
to predict the basis for immunogenic responses is also provided. The
methods can be used particularly for reducing the immunogenicity of
immunoglobulins or therapeutic proteins, e.g. Streptokinase (SK). The
products can be used for diagnosis and therapy. The present sequence
represents the amino acid sequence of the SK protein. (Updated on 17-OCT-
2003 to standardise OS field)
SQ Sequence 414 .AA;
Alignment Scores:
Pred. No.: 3,29e-176 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x AAW86143 (1-414)
QY 33 ATAGCTGGTCCGTAATGGCTACTAGATCGTCTTCGTAAATAACAGCCAAATTTGGTTGT 92
DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 93 AGCGTGTGCTGCTGCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40

QY 153 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACAGCAAGCTTAAAGTCCAAATCA 212
DB 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTTACTA 272
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 273 AAGCTATTCAAGACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTTCTTGAGTTC 332
DB 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 333 ATTGATTTTCAAGCGATGCAACATTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 392
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAAGATGGTTGCGTAACCTTTGCCGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACAT 452
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTCCGGTTAGACCATATAAAGAAAACCAATACAAAACCAAGGAAATCTGTTGATGTG 512
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGTAGTACGATTTTCACACAGGCTCTCAA 572
DB 161 GlutyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 632
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 633 GCTCAAGCAACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATCGTCACTCATGACATGATTTTCGTCAGATTTTACCAGTGGATCAAGAG 752
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTTACTTACCGTGTAAAAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
DB 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluLeuAsnLysLysSerGlyLeu 260
QY 813 AATGAAGAAATAACAACTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAGGG 872
DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 932
DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGTCGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAAACGCTAGCAACGTAAC 992
DB 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGACATTTATACGATCTCTGTCGTAAGGCTAAACTACTCTCAACAATCTC 1052
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1112
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAGATGCTAGCTAC 1172
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
DB 381 HisLeuAla 383

Search completed: January 28, 2006, 02:08:00
Job time : 126.195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 21.9832 Seconds
(without alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 2716
Sequence: 1 ttgttacttaagaaga.....ccttcacgatgtcgttag 1541

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xlP
-Q=/cgn2_1/USPTO_spool_p/US09940235/rnat_27012006_144219_27602/app_query.fasta_1.7708
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rat -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN 1.1.159 @rnat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/6_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/H_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/ECTUS_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/Backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984	73.0	413	2	US-10-360-101-264 Sequence 264, App
2	1984	73.0	414	2	US-09-211-542A-6 Sequence 6, Appl1
3	1984	73.0	414	6	5240845-1 Patent No. 5240845
4	1984	73.0	440	1	US-08-560-098A-52 Sequence 52, Appl1
5	1984	73.0	795	2	US-09-211-542A-2 Sequence 2, Appl1
6	1978	72.8	483	1	US-07-854-596B-47 Sequence 47, Appl1
7	1966.5	72.4	813	1	US-08-488-940-3 Sequence 3, Appl1
8	1966.5	72.4	1194	1	US-08-488-940-1 Sequence 1, Appl1
9	1963	72.3	435	1	US-07-854-596B-26 Sequence 26, Appl1
10	1958	72.1	440	1	US-07-854-596B-19 Sequence 19, Appl1
11	1958	72.1	440	1	US-07-854-596B-15 Sequence 15, Appl1
12	1958	72.1	483	1	US-07-854-596B-43 Sequence 43, Appl1

13	1958	72.1	499	1	US-07-854-596B-28 Sequence 28, Appl1
14	1958	72.1	859	1	US-07-854-596B-35 Sequence 35, Appl1
15	1937.5	71.3	1194	1	US-08-488-940-18 Sequence 18, Appl1
16	1936.5	71.3	1194	1	US-08-488-940-17 Sequence 17, Appl1
17	1928	71.0	413	1	US-08-759-559-12 Sequence 12, Appl1
18	1928	71.0	413	1	US-09-294-457-12 Sequence 12, Appl1
19	1928	71.0	413	2	US-09-919-703-12 Sequence 12, Appl1
20	1902	70.0	384	2	US-09-374-038-4 Sequence 4, Appl1
21	1902	70.0	384	2	US-09-658-179-4 Sequence 4, Appl1
22	1900	70.0	800	1	US-08-488-940-4 Sequence 4, Appl1
23	1900	70.0	1181	1	US-08-488-940-2 Sequence 2, Appl1
24	1898	69.9	372	2	US-09-374-038-3 Sequence 3, Appl1
25	1898	69.9	372	2	US-09-658-179-3 Sequence 3, Appl1
26	1890.5	69.6	747	1	US-07-854-596B-40 Sequence 40, Appl1
27	1887	69.5	401	2	US-09-374-038-1 Sequence 1, Appl1
28	1887	69.5	401	2	US-09-658-179-1 Sequence 1, Appl1
29	1887	69.5	413	2	US-09-374-038-2 Sequence 2, Appl1
30	1887	69.5	413	2	US-09-658-179-2 Sequence 2, Appl1
31	1875	69.0	369	1	US-07-854-596B-31 Sequence 31, Appl1
32	1695.5	62.4	736	2	US-09-211-542A-4 Sequence 4, Appl1
33	1686	62.1	356	2	US-09-211-542A-12 Sequence 12, Appl1
34	785	28.9	150	2	US-09-211-542A-14 Sequence 14, Appl1
35	632	23.3	2231	1	US-08-153-799-16 Sequence 16, Appl1
36	632	23.3	2324	1	US-08-283-857-1 Sequence 1, Appl1
37	632	23.3	2324	4	PCT-US95-09819-1 Sequence 1, Appl1
38	632	23.3	2355	2	US-10-360-101-235 Sequence 235, App
39	632	23.3	2386	1	US-09-016-366A-12 Sequence 12, Appl1
40	632	23.3	2386	2	US-09-961-403-1 Sequence 1, Appl1
41	632	23.3	2446	1	US-08-551-356-2 Sequence 2, Appl1
42	632	23.3	2446	4	PCT-US93-12687-2 Patent No. 5455158
43	627	23.1	2327	6	5455158-1 Sequence 4, Appl1
44	283	10.4	286	2	US-09-078-091-4 Sequence 2, Appl1
45	275	10.1	286	2	US-09-078-091-2 Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-10-360-101-264
; Sequence 264, Application US/10360101
; Patent No. 6861236
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: BP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:

Pred. No.:	2.88e-192	Length:	413
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x US-10-360-101-264 (1-413)

Qy	33	ATAGCGTCGCTGAAGGCTACGACGCTTCTTAATAACAGCAATGTTGTT	92
Db	1	11a1a1g1ypr0g1ut1r1p1e1u1a1p1a1r1p1r0s1e1a1b1a1n1s1e1r1g1n1e1u1a1v1a1	20


```

Db      ||| 81 LysAlaIleGlnIleuIleuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
Oy      ||| 333 ATTGATTTTGGCAAGCGCAACCATTAATCGATGGAAAGGCAAGCTCTTGGCTGAC 392
Db      ||| 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
Oy      ||| 393 AAGAGTGTTCGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGGTAAGCGGACAT 452
Db      ||| 121 LysAspGlySerValIleuProthInProValGlnIleuPheLeuSerGlyHis 140
Oy      ||| 453 GTGGCGGTAAAGACATATTAAGAAAAACAATACAAAACCAAGCGAAATCTGTGATGTG 512
Db      ||| 141 ValArgValArgProTyrIleuSgIlySproIleGlnAsnGlnAlaLysSerValAspVal 160
Oy      ||| 513 GAATATACGTGTACGTTTACCTCCCTTAAACCTGATGACAGATTTCAACAGGCTCAAA 572
Db      ||| 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Oy      ||| 573 GATACTAAGCTATTGAAACACTAGCTATCGTGACACATCACTCATCTCAAGAAATTA 632
Db      ||| 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu 200
Oy      ||| 633 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTTGAACGTGAC 692
Db      ||| 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArg 220
Oy      ||| 693 TCCTCAATGCTCACTATGACATGACATTTTCCGTAACGATTTTACCAATGATCAAGAG 752
Db      ||| 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspSgInGlu 240
Oy      ||| 753 TTACTTACCGTGTAAATGCGGAACAAGCTTATGATGCAATAAATCGTGTG 812
Db      ||| 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
Oy      ||| 813 AATGAGAATAAACAACAACCTGACCTGATCTGAGAAATATTACGCTTTAAAAAGGG 872
Db      ||| 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValIleuLysGly 280
Oy      ||| 873 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCACCATCAATACGTT 932
Db      ||| 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
Oy      ||| 933 GATTCGTATACCAACCAATTCCTTAAAGTGAAGACAGTCTTAAACGTTACGGAACGTAAC 992
Db      ||| 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 320
Oy      ||| 993 TTAGACTTCAGAGATTATACGATCCCTCGATAGAGCTTAACTACTCTACACAATCTC 1052
Db      ||| 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
Oy      ||| 1053 GATGCTTTTGGTATTATGACATATACCTTAACTGAAAGTAGAGTAATATCAAGATGAC 1112
Db      ||| 341 AspAlaPheGlyIleMetAspTyrThrIleuThrGlyLysValGlnAspAsnHisAspAsp 360
Oy      ||| 1113 ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAGAAGAGAAATGCTAGTAC 1172
Db      ||| 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAlaSerTyr 380
Oy      ||| 1173 CATTTAGCT 1181
Db      ||| 381 HisLeuAla 383

```

```

; FILING DATE: 06-JUL-1990
; SEQ ID NO: 1
; LENGTH: 414
5240845-1

Alignment Scores:
Pred. No.: 2,89e-192 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: Gaps: 6

US-09-940-235-9 (1-1541) x 5240845-1 (1-414)

Oy      ||| 33 ATAGCTGCTGTAATGCGCTACTAGATGCTCTTCTGTAAATTAACGCCAATGGTGT 92
Db      ||| 1 IleAlaGlyProGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
Oy      ||| 93 AGCGTGTGCTACTGTGTGAGGGGCAAGCAATTAAGCTTAAATTTTGAATC 152
Db      ||| 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
Oy      ||| 153 GATTAACATCAAGCACTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCAAAATCA 212
Db      ||| 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlnLysSerProLysSer 60
Oy      ||| 213 AAACATTTTGTCTAGTATGCGCGGATGTCACATTAATTTGAGAAAGCTGACTTACTA 272
Db      ||| 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysValAspLeuLeu 80
Oy      ||| 273 AAGGCTATTCAGAAACATTTGATGCTTAACCTGCAAGTCAAGCACTTACTTGAAGTGC 332
Db      ||| 81 LysAlaIleGlnGlnGlnIleuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
Oy      ||| 333 ATTGATTTTGGCAAGCGCAACCATTAATCGATGGAAAGGCAAGCTCTTGGCTGAC 392
Db      ||| 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
Oy      ||| 393 AAGAGTGTTCGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGGTAAGCGGACAT 452
Db      ||| 121 LysAspGlySerValIleuProthInProValGlnIleuPheLeuSerGlyHis 140
Oy      ||| 453 GTGGCGGTAAAGACATATTAAGAAAAACAATACAAAACCAAGCGAAATCTGTGATGTG 512
Db      ||| 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Oy      ||| 513 GAATATACGTGACGTTTACCTCCCTTAAACCTGATGACGATTTCAAGCGAGTCAAA 572
Db      ||| 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Oy      ||| 573 GATACTAAGCTATTGAAACACTAGCTATCGTGACACATCACTCATCTCAAGAAATTA 632
Db      ||| 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu 200
Oy      ||| 633 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCGGCTATACGATTTTGAACGTGAC 692
Db      ||| 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArg 220
Oy      ||| 693 TCCTCAATGCTCACTATGACCAATGACATTTTCCGTAACGATTTTACCAATGATCAAGAG 752
Db      ||| 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspSgInGlu 240
Oy      ||| 753 TTACTTACCGTGTAAATGCGGAACAAGCTTATGATGCAATAAATCGTGTG 812
Db      ||| 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
Oy      ||| 813 AATGAGAATAAACAACAACCTGACCTGATCTGAGAAATTTTGAAGTCTTAAAAAGGG 872
Db      ||| 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValIleuLysGly 280
Oy      ||| 873 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCACCATCAATACGTT 932

```

RESULT 3
5240845-1

; PATENT NO. 5240845
; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
; TSUTOMU

; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS

; NUMBER OF SEQUENCES: 65
; CURRENT APPLICATION DATA.

; APPLICATION NUMBER: US/07/549,049

Db 281 GluLysProTyrAspProHeaSpArgSerHisLeuLysLeuPheThrIleLeuTyrVal 300
QY 933 GATGTCGATACCAACGAATTGCTAAAGAGAGACAGCTCTTAACAGACTAGCGCAACGTAAAC 992
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluAsn 320
QY 993 TTAGACTTCAGAGATTTATATACGATCCTCGTGAATAGGCTAAACTACTTACCAACAACTTC 1052
Db 321 LeuAspPheArgAspLeuTyrAspProAlaGAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGCTATTATGAGCATATACCTTAAGTGAAGAGATTAATCAAGATGAC 1112
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACCGATACATAACGCTTATATAGGCAAGCAACCGAAGGAGAGAAATGCTACTAC 1172
Db 361 ThrAsnAlaGleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTAGCT 1181
Db 381 HisLeuAla 383

RESULT 4
US-08-560-098A-52
Sequence 52, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDEL, Stephan
APPLICANT: HEINZEL, WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSER: Evenson, McKee, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-52

Alignment Scores:
Pred. No.: 2,988-192 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0

DB: 1 Gaps: 0
US-09-940-235-9 (1-1541) x US-08-560-098A-52 (1-440)
QY 33 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTCTTGTAAATTAACAGCAATTGGTTGT 92
Db 27 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnSerGlnLeuVal 46
QY 93 AGCGTGGCTGCTCTGTGAGGGGAGCAATCAAGACACTTGTCTAAATTTTGAATC 152
Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspLysSerLeuLysPhePheGlnIle 66
QY 153 GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 212
Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
Db 213 AAACATTTGCTATCTGATAGTGGCGGAGATGTCATATACTTGAGAAAGCTGACTTACTA 272
Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 106
QY 273 AAGCTATTCAGAAACGAATTGATCGCTAACGTCACAGTAAACGACGACTTGAAGTC 332
Db 107 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 126
QY 333 ATGATTTTTCAGAGCATGCAACCATTAAGTATGATCGAAACGCAAGGCTACTTGTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAGATGGTTCGGTAACTTGGCCGACCCAACTGTCCAGAAATTTTGTAAAGCGACAT 452
Db 147 LysAspGlySerValThrLeuProGlnProValGlnGlnPheLeuLeuSerGlyHis 166
QY 453 GTGCGGCTTGAACCATTAAGAAACCAATCAAAACCAAGCAAGAAATCTGTGAAGTG 512
Db 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACGTGACAGTTTACTCCCTTAAACCTGATGACAGATTTCAGACAGGCTCAAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheAspProGlyLeuLys 206
QY 573 GATATTAAGCTATTGAAAACATTAAGTATGCGTGAACCATCATCATCAAGATTACTA 632
Db 207 AspThrLysLeuLeuLysThrIleValIleGlyAspThrIleThrSerGlnLeuLeu 226
QY 633 GCTGAAGCACAAGACATTTTAAACAAACCAACCCAGGCTATACGATTATGAACGTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATGCTCATGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTACTTACCGTGTAAATAATCCGGAACAAGCTTATAGATCAATAAATAATCGTCTG 812
Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaIleTyrArgIleAsnLysLysSerGlyLeu 286
QY 813 AATGAAGAAATAAACACATGACCTGATCTTGACAATATTAAGTCTTAAAAAGG 872
Db 287 AsnGluGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
QY 873 GAAAGCCGTATGATCCCTTGAATCGAGACATGAAACGTTCACATCAATTAAGTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCGATACCAACGAATTGCTAAAGAGAGACAGCTCTTAACAGCTAGCGCAACGTAAAC 992
Db 327 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluAsn 346
QY 993 TTAGACTTCAGAGATTTATATACGATCCTCGTGAATAGGCTAAACTACTTACCAACAACTTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProAlaGAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGCTATTATGAGCATATACCTTAAGTGAAGAGATTAATCAAGATGAC 1112

|||||
Db 367 AsphalaPheGlyIleMeLsPryrThrLeuThrgIySValGIuAspAsnHsAspAsp 386
Qy 1113 ACCAACCGTATCATTAACCGTTTATATGAGGACGCCGAGAGAGAAATGCTAGCTAC 1172
Db 387 ThrsAsnArgIleIleThrValIyrMeGlyLysArgProGIuGIuAsnAlaSerIyr 406
Qy 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
RESULT 5
US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-211-542A-2
Alignment Scores:
Pred. No.: 3,99e-192 Length: 795
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) * US-09-211-542A-2 (1-795)
Qy 33 ATAGCTGTCTGATAGTGGCTACTAGATCGTCTTCTGTAAATAACAGCAATGTTGTT 92
Db 382 IleAlaGlyProGIuIyrPheLeuLeuAspArgProSerValAsnAsnSerGIuLeuVal 401
Qy 93 AGCGTGTCTGTACTGTTCAGGGGAGCAATCAAGACTTAAATTTTGAATC 152
Db 402 SerValAlaGlyThrValGIuGIuIyrAsnGIuAspIleSerLeuIyrPheGluIle 421
Qy 153 GATCTAACGATCAAGACCTGCTCATGAGAGAAAGACAGAGCTTAAATCCAAATCA 212
|||||

Db 422 AspleuThrSerArgProAlaHisGIyGIyIyrThrgIuGIuGIuLeuSerProIySer 441
Qy 213 AACCATTTGGCTACTAGATGAGCGGAGATGTGACATAACTGAGAAAGCTGACTACTA 272
Db 442 LysProPheAlaIthrAspSerGIyAlaMetSerHisIySLeuGIuLysAlaAspleuLeu 461
Qy 273 AAGCTATTCAAGAACAAATTGATCGCTTAACGTCCACAGTAAACGACGACTACTTGAAGTTC 332
Db 462 LysAlaIleGIuGIuGIuLeuIleAlaAsnValHisSerAsnAspAspIyrPheGIuVal 481
Qy 333 ATTGATTTTGCAGAGGATGCAACCATTTACTGATCGAAACGGCAGAGTCTACTTCTGAC 392
Db 482 IleAspPheAlaSerAspAlaThrIleThrsArgAsnGIyIyValIyrPheAlaAsp 501
Qy 393 AAAGATGCTGGTGAACCTTGCCGACCCGACCTGTGCCAAGAAATTTTGCTAAGCGGACAT 452
Db 502 LysAspGIySerValThrLeuProThrGIuProValGIuGIuPheLeuSerGIyHis 521
Qy 453 GTGGCGCTTAGACCATATATAAGAAACCAATATCAAAACCAAGCGAAATCTGTGATGTG 512
Db 522 ValArgValArgProIyrIyGIuIySerProIleGIuAsnGIuAlaIySerValAspVal 541
Qy 513 GAATATACTGTACAGTTACTCTCCTTAACCCCTGATGACATTTGACACAGTCTCAAA 572
Db 542 GIuIyrThrValGIuPheThrProLeuAsnProAspAspPheArgProGIyLeuIyS 561
Qy 573 GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATTCACATCTTCAGAAATTACTA 632
Db 562 AspThrIySLeuLeuIySThrLeuAlaIleGIyAspThrIleThrsGIuGIuLeuLeu 581
Qy 633 GCTCAAGACAAAGCATTTTAAACAAAACCAACCCAGGCTTACGATTTTATGACAGTAC 692
Db 582 AlaGIuAlaGlnSerIleLeuAsnLysAsnHisProGIyIyrThrIleIyGIuArgAsp 601
Qy 693 TCCTCAATCGTCACTGACATGACAAATGACATTTTCGTGATTTTCAATGATCAAG 752
Db 602 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGIuIu 621
Qy 753 TTTACTTACCGTGTAAAAATCGGAGACAGCTTATAGATCATATAAAAAATCGTGTG 812
Db 622 PheThrIyrArgValIySAsnArgGIuGIuAlaIyrArgIleAsnLysIySerGIyLeu 641
Qy 813 AATGAGAAATTAACAAACACGACCTGATCTGTAGAAATTTACGCTTAAAAAGG 872
Db 642 AsnGIuGIuIleAsnAsnThrAspLeuIleSerGIuIySyrIyValLeuLysIyGIy 661
Qy 873 GAAAGCCGATGTGATCCCTTGTATCGCAGTCACTTGAACCTTCAACATCAATACGTT 932
Db 662 GIuIySProIyrAspProPheAspArgSerHisLeuIySLeuPheThrIleIyGIyVal 681
Qy 933 GATGTGATACCAACGAATTGCTTAAAAAGTAGAGCTCTTAAACGCTAGCAAGTAC 992
Db 682 AspValAspThrAsnGIuLeuLeuIySerGIuGIuLeuThrAlaSerGIuArgAsn 701
Qy 993 TTACACTTCAGAGATTTTATAGATCCCTGTGATAGGTAAGCTTAACTTACAACAATTC 1052
Db 702 LeuAspPheArgAspLeuIyrAspProArgAspLysAlaIySLeuIyThrAspAsnLeu 721
Qy 1053 GATGCTTTGGTATTATGACTATACCTTACGTGAAAGTAGAGATTAACAGATAC 1112
Db 722 AspAlaPheGIyIleMetAspIyrThrLeuThrgIySValGIuAspAsnHisAspAsp 741
Qy 1113 ACCAACCGTATCATTAACCGTTTATATGAGGACGCCGAGAGAGAAATGCTAGCTAC 1172
Db 742 ThrsAsnArgIleIleThrValIyrMeGIyLysArgProGIuGIuAsnAlaSerIyr 761
Qy 1173 CATTTAGCT 1181
Db 762 HisLeuAla 764
RESULT 6
US-07-854-596B-47
; Sequence 47, Application US/07854596B

```

: Patent No. 5434073
: GENERAL INFORMATION:
: APPLICANT: Dawson, Kelch M
: APPLICANT: Hunter, Michael G
: APPLICANT: Caplewski, Lloyd G
: TITLE OF INVENTION: Proteins and nucleic acids
: NUMBER OF SEQUENCES: 73
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dr. John J. McDonnell
: STREET: Ten South Wacker Drive, Suite 3000
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/854,596B
: FILING DATE: 03-JUN-1992
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: McDonnell, John J
: REGISTRATION NUMBER: 26,949
: REFERENCE/DOCKET NUMBER: 92,337
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-715-1000
: TELEFAX: 312-715-1234
: TELEX: 910-221-5317
: INFORMATION FOR SEQ ID NO: 47:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 483 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-854-596B-47

```

```

Alignment Scores:
Pred. No.: 1,266-191 Length: 483
Score: 1978.00 Matches: 401
Percent Similarity: 82.28% Conservative: 17
Best Local Similarity: 78.94% Mismatches: 52
Query Match: 72.83% Indels: 38
DB: 1 Gaps: 8

```

US-09-940-235-9 (1-1541) x US-07-854-596B-47 (1-483)

```

QY 33 ATAGTGTCTGATGATGCTACTAGATCGCTTCTGTAATAATACAGCAATTGTTGTT 92
Db 1 IleaaglyProglutripLeuaspaspProSerValasnaserGlnleuVal 20
QY 93 AGCGTGTGTAAGTGTGAGGAGGAGCAATCAAGCATTAATTTTGAATC 152
Db 21 SerValaaglyThnValGlnGlyThranglnaSerleuyspheeGlnVal 40
QY 153 GATTCATCATCAAGCCTGCTCATGAGAGAAAGACAGCAAGCATTAAGTCCAAATCA 212
Db 41 AspleuThrSerArgProAlaHieGlyGlystnGlnGlnGlyleuSerProlysSer 60
QY 213 AAACCATTTGCTACTGATAGTGGCGGAGTGTCACTAACTTGAGAACTGACTACTA 272
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHslsyleuGlnLysAlaAspleuLeu 80
QY 273 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAGAGAGCATTTGAGTGC 332
Db 81 LysAlaIleGlnGlnGlnleuIleAlaasnValHsSerasnspasyrPheeGlnVal 100
QY 333 ATTGAATTTGCAAGGATGCAACCATTACTGATCGAAGCGCAAGGCTTACTTTGCTGAC 392
Db 101 IleaspPheAlaSerAspAlaThrIleThrAspArgasnGlyLysValIlyrPheAlaasp 120

```

```

QY 393 AAAGATGTCGGTAAACCTTGCCGAGCCCAACCTGTCACAGAAATTTTGTAAAGCGACAT 452
Db 121 LysaspGlySerValThrleuProThnGlnProValGlnGlnleuLeuSerGlyHis 140
QY 453 GTGCGGCTTAGACCATTAATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTC 512
Db 141 ValArgValArgProGlyLysGlnLysProIleGlnasnGlnAlaLysSerValAspVal 160
QY 513 GAATATATCTGATAGTTTACTCCCTTAACCCCTGAAGTACGATTTTCAGACAGGCTGAAA 572
Db 161 GlnIyThrValGlnPheThrProleuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATATTAAGCTATTGAAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632
Db 181 AspThrLysleuLeuysThrleuAlaIleGlyAspThrIleThrSerGlnleuLeu 200
QY 633 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTTACGATTTATGAACGTGAC 692
Db 201 AlaglnAlaGlnSerIleleuAsnLysThrHisProGlyIyThrIleIyGlnArgAsp 220
QY 693 TCCCTCAATCGTCACTGACATGACATGACATTTCCGTCGATTTTACCAATGATCAAGAG 752
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleleuPrometAspGlnGln 240
QY 753 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTAATGATCAATAAAAAATCGTCTG 812
Db 241 PheThrIyHisValIlyssasnArgGlnGlnAlaIyGlnIleasnLysLysSerGlyLeu 260
QY 813 AATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTACGTCCTTAAAAAGG 872
Db 261 AsnGlnGlnIleasnThrAspIleSerGlnLysIyThrIyValleuLysLysGly 280
QY 873 GAAAAGCCGTAATGATCCCTTGTATCGGAGTCACTTGAACCTGTACCATCAATATGCTT 932
Db 281 GlnLysProIyAspProPheAspArgSerHisleuLysleuPheThrIleLysIyVal 300
QY 933 GATGTCGATCAACAGCAATGCTTAAAGAGACAGCTTTAACAGCTAGCGAAGCTGAC 992
Db 301 AspValAsnThrAsnGlnleuLeuLysSerGlnGlnleuThrAlaSerGlnArgAsn 320
QY 993 TTAGACTTCAGAGATTATATCGATCCCTCGATATAGGCTTAACTCTACCAACATCTC 1052
Db 321 LeuAspPheArgspLeuIyAspProArgAspLysAlaLysleuLeuIyAsnAsnleu 340
QY 1053 GATGCTTTTGGTATATGACATTAACCTTAACCTGAAAAGTAAATTAACAGATGAC 1112
Db 341 AspAlaPheGlyIleleuAspIyThrleuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1113 ACCAACCTATCATTAACCGTTTATATGCGGAAGCAACCGAAGAGAAATGCTAGCTAC 1172
Db 361 ThrAsnArgIleIleThrValIyMetGlyLysArgProGlnGlnGlnAlaSerIy 380
QY 1173 CATTTAGCTGTGTGTGCGCCAGCGCAAGATTTGACCCATAGCTGAAAGTGTTTGAT 1232
Db 381 HisleuAla-----IyAspLysAspArgIyThrGlnGlnGlnAlaGlnValIySer 398
QY 1233 CATGCTGTGGAAGCTTCTATATGTCGCGAAGAACTGGAGAAACCTTACCAAGCTGG 1292
Db 399 TyrleuArgIyThrGlnIyThrProIleProAspAsnProAsnAspLysIleGlnGlyArg 418
QY 1293 ATGATG---CTAATTTGACT-----TGCCTGGAGAAAGGAGGGA 1331
Db 419 ValValIyThrAspCysThrGlnSerGlyGlnAsnleuLysCysGlnGlySerAsn 438
QY 1332 CGCATCACTTGCACTTCTAAGAAATTAATGTCAGATCAGACACAAAGATCCTATAGA 1391
Db 439 -----ValCysGlnGlnGlnLysAsnLysCys-----Ile 447
QY 1392 ATTGAAGACACTTGAGCAAGAAAGATTAATGAGAAACCTGCTCAGTGCATCTGACA 1451
Db 448 LeuGlySeraspGlyGlnLysAsn-----GlnCysVal---Thr 459
QY 1452 GCGAAGCGCCGAGAGAGTGAAGTGTGAGAGGACACACTCTGTGACAGACCAATCGAC 1511

```

Db 460 GlyGluGly-----ThrProLysProGlnSerHis 469
QY 1512 GGATCTGGCCCTTCAACCATGTT 1535
Db 470 AsnAspGlyAspPheGluGluIle 477
RESULT 7
US-08-488-940-3
; Sequence 3, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 813 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-940-3
Alignment Scores:
Pred. No.: 2,4e-190 Length: 813
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 72.40% Indels: 1
DB: 1 Gaps: 1
US-09-940-235-9 (1-1541) x US-08-488-940-3 (1-813)
QY 33 ATAGCTGCTCTGAATGCTACTAGATCGCTTCTGTAATAAAGCCAAATGTTGTT 92
Db 1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 93 AGCGTGTCTGTAAGTCTTGAAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 152
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
QY 153 GATCTTAACATCAAGACCTGCTCATAGAGAGAAAGACAGAGAGCTTAAGTCCAAATCA 212
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCAATTAACCTTGAAGAGCTGACTACTA 272
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80

QY 273 AAGCTATTCAAGAAACATTGATGCTTAAGCTCCACAGTAAAGCACTACTTTGAGGTC 332
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspArgPheGluVal 100
QY 333 ATTGATTTTGGAGCGATGCAACCATTAATGATCGAAAACGGCAAGGTCTTGTCTGAC 392
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAAGATGTTGGGTAACTTGGCCGACCCACCTGTCCCAAGAAATTTTGTCTAAGGAGAT 452
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGGCGCTTAGACCATTTAAAGAAAACCAATACAAAACCAAGCGAAATCTGTGATGTC 512
Db 141 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
QY 513 GAATATACCTGATACGTTACTCCCTTAAACCTGATGACGATTTCAACACAGTCTCAAA 572
Db 160 GluThrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
QY 573 GATCTTAAGCTATTGAAAACATGCTATCGGTGACACATCACTCAATCAAGAAATTACTA 632
Db 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 199
QY 633 GCTCAAGCAACAAACATTTTAAACAAAACACACCGGCTTATGATTTAAGACGTGAC 692
Db 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
QY 693 TCCTCAATCGTCACTCACTGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAG 752
Db 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
QY 753 TTACTTACCGGTGTTAAAAATCGGAAACAGCTTATAGATCAATATAAAAATCTGCTCG 812
Db 240 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysSerGlyLeu 259
QY 813 AATCAAGAAATTAACAACACCTGACCTGATCTGTGAAATTTACGTCCTTAAAAAGCG 872
Db 260 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 279
QY 873 GAAAAGCGTATGATCCCTTGAATCGGACGATCACTTGAACCTTACCATCAATGACTC 932
Db 280 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 299
QY 933 GATGTGATACCAACGAATGCTTAAAAAGTAGAGAGCTTTAACAGCTAGGAACTGAC 992
Db 300 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 319
QY 993 TTAGACTTCAGAGATTTATACGATCTCGTATAGGCTTAACCTTACTTACACAAATCTC 1052
Db 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAspAsnLeu 339
QY 1053 GATGCTTGTGATTAATGACCTATACCTTAACTGAAAAGTAGAGAGATTAACAGATGAC 1112
Db 340 AspAlaPheGlyIleLeuAspTyrThrLeuThrGlyValGluAspAsnHisAspAsp 359
QY 1113 ACCAACCGTATCAATACCGTTTATATAGGCAAGGACCCGAAAGAGAGAAATGTAGTAC 1172
Db 360 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 379
QY 1173 CATTTAGCT 1181
Db 380 HisLeuAla 382
RESULT 8
US-08-488-940-1
; Sequence 1, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20

```

CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,940
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Frazer, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-940-1

Alignment Scores:
Pred. No.: 2.9e-190 Length: 1194
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 72.40% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-9 (1-1541) x US-08-488-940-1 (1-1194)
QY 33 ATAGTGGTCTCGATGCTGCTACTAGATGCTCTTCTGTAATTAACGCCAATTGTTGTT 92
DB 382 TleaaaglyprrglttrprrleuunspargprrserValasnaaserrglnleuVala 401
QY 93 AGCGTGGTGGTACTGTTGAGGGGAGCGAATCAAGACATTAGTCTTAATTTTGAATC 152
DB 402 SerValaIaaglyThrValaIaaglyThraaaglnaasprrSerleuulysPhepnegluile 421
QY 153 GATCAACATCAAGCACTGCTCATGAGGAGAAAGACAGACAAAGCTTAAGTCCAAATCA 212
DB 422 AspleuThrSerArgPrrroahisgilylYThrGlnGlnGlyleuSerProlysser 441
QY 213 AAACCATTTGCTACTGATGATGTCGGCGGAGTGCATAACTTGAGAAAGCTGACTTACTA 272
DB 442 LysPrrPrrlealathraasprerGlyAlaMetSerHislyleuGlnIulysaIaaspleu 461
QY 273 AAGGCTATTCAGAGCAATTCATGCTAAGTCCACAGTACAGAGCACTTATGAGTTC 332
DB 462 LysaIaIaIeGlnGlnGlnleuIlealaaanValHisSerasnaasprpyrPhegluVal 481
QY 333 ATTGATTTTGCAGAGCATGACCAACTTACTGATCGAAGCGCAAGGCTTACTTGTGAC 392
DB 482 IleasprPhealaserAspralathrIleThraasprasngilylYsValYrrPhealaa 501
QY 393 AAAGATGGTTCGGTAACTTGTCCGACCAACTGTCACAAAGATTTTGTCAAGCGACAT 452
DB 502 LysasprGlySerValThrleuPrrThrGlnProValaGlnGlnIupheleuSerGlyHis 521
QY 453 GTGGCGGTAGACCATATTAAGAAAAACAATTCAAACCAAGCGCAATCTGTGTATGTG 512

```

```

DB 522 ValArgValArg--TyrIysGlnIysProIleGlnaaglnaIalYsSerValaAspVal 540
QY 513 GAATATACGTAGACTTACTCCCTTAACCGTAGTACGATTTTCAGACCAAGCTCAAA 572
DB 541 GlutYThrValaGlnPrrThrProleuAsnProasprasprrPheApgProGlyleu 560
QY 573 GATCACTAAGCTATTTGAAAAACATAGCTATGTCGACACATCATCATCTCAAGATTTACTA 632
DB 561 AsprThrIysleuIeulysThrIleuAlaIleGlyAsprThrIleThrSerGlnIleu 580
QY 633 GCTCAAGCAACAAGCATTTTAAACAAAACCAACCGACGCTATACGATTTTAAAGCGTAC 692
DB 581 AlaGlnIaGlnIleSerIleleuAsnIysAsnHisProGlyYrrThrIleYrrGlnArg 600
QY 693 TCCTCAATGCTCATGACATGACATTTTCCGTAACGATTTTACCAAGGATCAAGAG 752
DB 601 SerSerIleValThrHisasprAsnAsprrIlePheArgThrIleuPrrMetAspGln 620
QY 753 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATGATCAATAAAAAATCGTCTG 812
DB 621 PrrThrYrrArgValaIlysaasnaargGlnGlnAlaIrrArgIleasnylysserGly 640
QY 813 AATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTACGTCCTTAAAAAAGG 872
DB 641 AsnGlnGlnIleleuAsnThrAsprleuIleSerGlnIysYrrYrrValleuIysGly 660
QY 873 GAAAAGCGGTATGATCCCTTGTATCGCAGTCACTTGAACCTGTCACATCAATACGTT 932
DB 661 GlulysPrrGlyAsprPrrPrrleuAsprSerHisIleuIysleuPrrThrIleYrrVal 680
QY 933 GATGTGATACCAAGCAATGTGTTAAAAAGGACAGCTCTTAACAGCTACGGAACGTAAC 992
DB 681 AsprValaAsprThrAsnGlnleuIeulysSerGlnGlnleuIleThrAlaSerGlnArg 700
QY 993 TTGACTTCAGAGATTATTCAGATCTCGTATAGGCTTAACACTACTCAACAACTTC 1052
DB 701 leuAsprPheargAsprleuYrrAsprPrrroahisgilylYsleuIeulYrrAsna 720
QY 1053 GATGCTTTGGTATTTAGACTATACCTTAACCTGAGAAAGTGAAGATTAATCACGATGAC 1112
DB 721 AsprAlaPrrleGlyIleleuAsprYrrThrleuThrGlyYsValaGlnAsprAsnHis 740
QY 1113 ACCAACCTTACATTAACCGTTTATATGAGGACGACCGAAGAGGAATGCTAGTAC 1172
DB 741 ThrAsnArgIleIleThrValaIrrMetGlyYsArgPrrroGlnGlnIulaasnaIa 760
QY 1173 CATTAGCT 1181
DB 761 HisleuAla 763

RESULT 9
US-07-854-596B-26
; Sequence 26, Application US/07854596B
; Patent No. 5434073
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Czaplowski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSER: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDowell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-596B-26

Alignment Scores:
Pred. No.: 3,89e-190 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 72.28% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-9 (1-1541) x US-07-854-596B-26 (1-415)

QY 30 ATGATAGCTGCTCTGTAAGTGGCTACTAGATGCTCTTGTGTAATTAACGCAATTGGTT 89
DB 1 MetIleAglYProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 90 GTTACGCTTGTGCTACTGTTGAGGGGAGCATACAGACATTAGCTTAAATTTTGTGA 149
DB 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGlu 40
QY 150 ATCGATCTAACATCAACGACTGCTCATGAGAGAAAGACAGCAAGGCTTAGTCCAAA 209
DB 41 IleAspLeuThrSerArgProAlaHisGlyGlyLysTrpGluGlnGlyLeuSerProLys 60
QY 210 TCAGAACCATTTGCTACTGATAGTGGCGGATGTCACATTAACCTTAGAAGCTGACTTA 269
DB 61 SerArgProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 80
QY 270 CTAAAGGCTATTCAAGAACATTGATCGCTAACGTCACAGTACAGCAGCTACTTGA 329
DB 81 LeuLysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgPheGlu 100
QY 330 GTGATGATTTTGAAGGATGACCAACATTACTGATGAAACGGCAAGGCTACTTGT 389
DB 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlyrPheAla 120
QY 390 GACAAAGATGGTGGTAACTTCCGACCCGACCACTGTCGCAAGAAATTTTGTAGCGGA 449
DB 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGly 140
QY 450 CATTCGCGCTTAGACCATATTAAGAAAACCAATATCAAAAACAAGCAAGAAATCTGTAT 509
DB 141 HisValArgValAlaGlyProIlyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAsp 160
QY 510 GTGGAATATCTGACGATTACTCCCTTAAACCTGATGACGATTTCAGACAGGCTTC 569
DB 161 ValGlnLysThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeu 180
QY 570 AAGATCTAGCAATTTGAAGAACTAGCTATCGGTGACACATACATCTCAAGATTA 629
DB 181 LysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeu 200
QY 630 CTAGCTCAAGCAACAGATTTTAAACAAAACCAACGAGCTATACGATTTATGAAGT 689
DB 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyIlyrThrIleIlyrGlnArg 220

QY 690 GACTCTCAATCGCTACCTGACATGACAAATGACATTTCCGTAAGATTTTACCAATGATCA 749
DB 221 AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAlaGln 240
QY 750 GAGTTTACTTACCGTGTGTTAAATCCGGAAACAGCTTATAGATCATTAATAATCTGCT 809
DB 241 GluPheThrIlyrHisValLysAsnArgGlnGlnAlaIlyrGlnIleAsnLysSerGly 260
QY 810 CTGAATGAAGAAATTAACCAACACTGACCTGATCTGAGAAATTTAGTCCTTAAAAA 869
DB 261 LeuAsnGlnGlnIleAsnThrAsnThrAspLeuIleSerGlnLysIlyrIlyrValLeuLys 280
QY 870 GGGGAAAGCCGTATGATCCCTTGAATCGACATCACTTGAACCTTCCACCATCAATAC 929
DB 281 GlyLysLeuProIlyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIlyr 300
QY 930 GTTGAATGCTGATCCCAACGAAATTCCTTAAAGTAGAGAGCTTTAACAGCTAGCAAGT 989
DB 301 ValAspValAlaSerThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArg 320
QY 990 AACTAGACTTCAGAGATTTTATACGATCCCTCGATTAAGGCTAACTTCAACAAAT 1049
DB 321 AsnLeuAspPheArgAspLeuIlyrAspProArgAspLysAlaLysLeuLeuIlyrAsnAsn 340
QY 1050 CTCGATGCTTTTGTATTATGACTATACCTTAACCTGGAAGTAGAGATTAATCACGAT 1109
DB 341 LeuAspAlaPheGlyIleMetAspIlyrThrLeuThrGlyLysValGlnAspAsnHisAsp 360
QY 1110 GACACCAACCGTATCATTAACCGTTTATATGAGGACAGCAGGAGAGAGAAATGCTAGC 1169
DB 361 AspThrAsnArgIleIleThrValIlyrMetGlyLysArgProGlnGlyGlnAsnAlaSer 380
QY 1170 TACGATTTAGCT 1181
DB 381 TyrHisLeuAla 384

RESULT 10

US-07-854-596B-19

/ Sequence 19, Application US/07854596B

/ Patent No. 5434073

/ GENERAL INFORMATION:

/ APPLICANT: Dawson, Keith M

/ APPLICANT: Hunter, Michael G

/ APPLICANT: Czaplowski, Lloyd G

/ TITLE OF INVENTION: Proteins and nucleic acids

/ NUMBER OF SEQUENCES: 73

/ CORRESPONDENCE ADDRESSES:

/ ADDRESSER: Dr. John J. McDowell

/ STREET: Ten South Wacker Drive, Suite 3000

/ CITY: Chicago

/ STATE: IL

/ COUNTRY: USA

/ ZIP: 60606

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: IBM PC compatible

/ SOFTWARE: Patentin Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/07/854,596B

/ FILING DATE: 03-JUN-1992

/ CLASSIFICATION: 435

/ ATTORNEY/AGENT INFORMATION:

/ NAME: McDowell, John J

/ REGISTRATION NUMBER: 26,949

/ REFERENCE/DOCKET NUMBER: 92,337

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 312-715-1234

/ TELEX: 910-221-5317

/ INFORMATION FOR SEQ ID NO: 19:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 435 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-596B-19

Alignment Scores:

Pred. No.:	1.28e-189	Length:	435
Score:	1958.00	Matches:	378
Percent Similarity:	98.96%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	4
Query Match:	72.09%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-9 (1-1541) x US-07-854-596B-19 (1-435)

```

QY 33 ATAGTGTGCTCCGATGAGTGGTACTAGATCGTCTCTGTGTAATTAACGCCAATGGTGTGTT 92
Db 22 IleaGlyProGluThrLeuAspArgProSerValAsnAsnSerGlnLeuValVal 41
QY 93 AGCGTTGCTGTAAGTGGGGAAGATCAAGACATTAAGTCTTAATTTTGAATC 152
Db 42 SerValaGlyThrValGlnGlnThrAsnGlnAspLysSerLeuLysPheGlnIle 61
QY 153 GATCAATCAACGACCTGCTCATGAGGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
Db 62 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 81
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAATTGAGAAAGCTGACTTA 272
Db 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 101
QY 273 AAGGTATTCAAGAACATTTGATCGCTAACGTCACAGTAAGACGACTTGTAGGTC 332
Db 102 LysAlaIleGlnGlnGlnLeuIleAsnValHisSerAsnAspArgPyrPheGluVal 121
QY 333 ATTGATTTTGGCAAGGAGATGCAACCATTAATCTGATCCGAAAGGCAAGGCTTGTGTCAC 392
Db 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValValIyrPheAlaAsp 141
QY 393 AAAGTGGTTCGTAACTTGGCCGACCCCAACCTGTCCAGAAATTTTGTAAAGCGGACAT 452
Db 142 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 161
QY 453 GTGGCGCTTACACCATTAATTAAGAAAAACCAATACAAAACCAAGCGAAATCTGTATGTCG 512
Db 162 ValArgValArgProGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 181
QY 513 GAATATCTGTACAGTTACTCCCTTAAACCTGATGACGATTTTTCAGACAGGTCGAA 572
Db 182 GluThrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
QY 573 GATACATAAGCTAATTAAGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 632
Db 202 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 221
QY 633 GCTCAAGCAACAAGCATTTTAACAAAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyLysThrIleLysGlnLysGly 241
QY 693 TCTCAATCGTCACTCATGACATGACATTTCCGTAAGATTTTTCAGATGATGACAGAG 752
Db 242 SerSerIleValThrHisAspAsnAspIlePheAsnGlnThrIleLeuProMetAspGlnGln 261
QY 753 TTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATTAATAATCTGGTCTG 812
Db 262 PheThrThrHisValLysAsnArgGlnGlnAlaThrGlnIleAsnLysLysSerGlyLeu 281
QY 813 AATGAAGAAATTAACAACACTGATCTGTGAGAAATATTACGCTCTTAATAAAGG 872
Db 282 AsnGlnGlnLysLeuAsnThrAspLeuLysSerGlnLysLysLysValLeuLysLysGly 301
QY 873 GAAAGCGGTATGATCCCTTATGATGACGATCACTTGAACATGCTTACCAATCAATCGCT 932

```

```

Db 302 GluLysProLysAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
QY 933 GATGCGATTAACCAATTTGCTAAAGAGACAGCTCTTAACAGTACGAAACGTAAC 992
Db 322 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysGln 341
QY 993 TTAGCTCGAGAGATTTATACGATCCGCTGATAGGCTAAAGTACTCTCAACAATCTG 1052
Db 342 LeuAspPheArgAspLeuThrAspProLysAspLysAlaLysLeuLeuThrAsnAsnLeu 361
QY 1053 GATGCTTTTGGATTAATGACATTAACCTTAAGTGAAGAAAGTAAAGATTAATCAAGTAC 1112
Db 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlnLysValGlnAspAsnHisAspAsp 381
QY 1113 ACCAAGCTATCAATACCGTTTATATGGGACAGACCCGAAAGAGAAATGCTAGTAC 1172
Db 382 ThrAsnArgIleLeuThrValThrMetGlyLysArgProGlnGlnLysAlaSerTyr 401
QY 1173 CATTTAGCT 1181
Db 402 HisLeuAla 404

RESULT 11
US-07-854-596B-15
; Sequence 15, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-15

Alignment Scores:
Pred. No.: 1.29e-189 Length: 440
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-9 (1-1541) x US-07-854-596B-15 (1-440)

```



```

QY 33 ATAGCTGCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGGTGT 92
DB 27 ILeAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 93 AACGTTGCTGTACTGTGGAGGAGAAATCAAGACATTAGTCTTAAATTTTGAATC 152
DB 47 SerValAlaGlyThrValGluGlnAsnGlnAspIleSerLeuLysPhePheGluIle 66
QY 153 GATCTAACATCAGCACTGCTCATGAGAGAAAGAGAGAGAGGCTTAAGTCCAAATCA 212
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlnLeuSerProLysSer 86
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 272
DB 87 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysValAspLeuLeu 106
QY 273 AAGGCTATTCAGAGAAACATTTGATCGCTAACGTCCAGTACGACGACTTCTGAGGTC 332
DB 107 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGlnVal 126
QY 333 ATTGATTTTGAAGCGATGCAACCATTAAGTATGCAAGCGGAGGCTTACTTGTCTGAC 392
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAGATGCTTCGGTAACTTTGCCGACCCAACTGTCGCAAGATTTTGTGAAGCGGACAT 452
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnLeuLeuSerGlyHis 166
QY 453 GTGCGGCTTAAACCATATATAAGAAAAACAATACAAACCAAGCGAAATGTTGATGTG 512
DB 167 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACGAGCTTCAA 572
DB 187 GlnTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATCTAAGCTATTTGAAAAACATAGCTATCGGTGACACCATGCAATCTCAAGAAATTA 632
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 633 GCTCAGACCAAGCAATTTTAAACAACAAACCAAGGCTATACGATTTTGAACGCTGAC 692
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnLysArg 246
QY 693 TCTTCAATGCTACATCATGACATGACATTTTCCGATCGAATTTTCAATGATGATCAAG 752
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGGTAAAAAATGGGAAACAAGCTTATAGATCAATAAAAAATCTGCTCG 812
DB 267 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnLysLysSerGlyLeu 286
QY 813 AATGAAGAAATTAACAACAACATGACTGATCTCTGAGAAATATATACGCTTAAAAAAAGG 872
DB 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 306
QY 873 GAAAAAGCGTATGATCCCTTGTATGCGACGACTTGAACCTGTTCACATCAAAATACGT 932
DB 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATTCGCAATACCAAGAAATTTGCTAAAAAATGAGACGACTTTAACAGCTACGAAAGTAA 992
DB 327 AspValAlaGlyThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 346
QY 993 TTAGACTTCAGAGATTTATACGATCTCGTGAATAAGGCTAACTACTTCAACAATTC 1052
DB 347 LeuAspPheAspAspLeuTyrAspProAspAspLysAlaLysLeuLysTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGTATATGACATATACCTTAACTGGAAGAGAGATATACGATAC 1112
DB 367 AspAlaPheGlyIleMetAspTyrThrIleuThrGlnLysValGlnLysAsnHisAspAsp 386

```

```

QY 1113 ACCAAGCTATCATTAACCGTTTATATGGGCAAGGACCCGAAAGAGAGAAATGCTAGTAC 1172
DB 387 ThrAsnArgIleIleThrValTyrMetCglyAsnArgProGlnGlnLysAlaSerTyr 406
QY 1173 CATTTAGCT 1181
DB 407 HisLeuAla 409

RESULT 12
US-07-854-596B-43
; Sequence 43, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SRO ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-43

Alignment Scores:
Pred. No.: 1,35e-189 Length: 483
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-9 (1-1541) x US-07-854-596B-43 (1-483)
QY 33 ATAGCTGCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGGTGT 92
DB 27 ILeAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 89
QY 93 AACGTTGCTGTACTGTGGAGGAGAAATCAAGACATTAGTCTTAAATTTTGAATC 152
DB 47 SerValAlaGlyThrValGluGlnAsnGlnAspIleSerLeuLysPhePheGluIle 109
QY 153 GATCTAACATCAGCACTGCTCATGAGAGAAAGAGAGAGAGGCTTAAGTCCAAATCA 212
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlnLeuSerProLysSer 129
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 272

```

```

Db      ||| 130 LysProPheAlaThrAspSerGlyAlaMetProHisLeuGluValysAlaAspLeuLeu 149
Qy      ||| 273 AAGGTAATTCAGAGCAATTCGCTAAGCGCCAGATGAACAAGCACTTGTGAGTC 332
Db      ||| 150 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 169
Qy      ||| 333 ATTGATTTTGCAGCGATGCAACCATTCATGATCGAAAGCGCAAGGCTCTTGTGTCAC 392
Db      ||| 170 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 189
Qy      ||| 393 AAAAGATGGTTGGGTAACTTGGCCGAGCCAACTGTCGAAGATTTTGTCAAGCGACAT 452
Db      ||| 190 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 209
Qy      ||| 453 GTGCGCGGTGAGCACTATTAAGAAAAACAATACAAACCAAGCGAAATCTGTGATGTCG 512
Db      ||| 210 ValArgValArgProTyrGlyGlyProIleGlnAsnGlnAlaLysSerValAspVal 229
Qy      ||| 513 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTGAGACGAGGTCGAA 572
Db      ||| 230 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 249
Qy      ||| 573 GATACCTAAGCTATTGAAAACTAGCTATCGGTGACACCATGACATCTCAAGATTAATA 632
Db      ||| 250 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 269
Qy      ||| 633 GCTCAAGCAACAAGCAATTTTAAACAACAAACCAAGCGCTATACGATTTATGACGTCAC 692
Db      ||| 270 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 289
Qy      ||| 693 TCCTCAATCGTCACTCATGACATGACATTTCCGATACGATTTTACCATGATGCAAGAG 752
Db      ||| 290 SerSerIleValAlaThrHisAspAsnAspIlePheAspGlyThrIleLeuProPheAspGlnGln 309
Qy      ||| 753 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATATAAATCGTGTG 812
Db      ||| 310 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnLysLysSerGlyLeu 329
Qy      ||| 813 AATGAAGAATATTAACAACACTGACCTGATCTCTGAGAAATATTAAGCTCTTAAAAAGAG 872
Db      ||| 330 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 349
Qy      ||| 873 GAAAGCCGTAATGATCCCTTGTGATGCGCAATCACTGAACTGTTCACATCAATCGTT 932
Db      ||| 350 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 369
Qy      ||| 933 GATGTCGATACCAAGCAATTCGCTAATAAAGTACAGCTCTTAAACAGCTACGCAAGTAAAC 992
Db      ||| 370 AspValAlaThrHisGlnLeuLeuLysSerGlnGlnIleLeuThrAlaSerGlnAlaGln 389
Qy      ||| 993 TTAGACTTCAGAGATTTATACGATCTCGTATTAAGGCTTAAACTACTTACCAAGATCTTC 1052
Db      ||| 390 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 409
Qy      ||| 1053 GATGCTTTTGTGATTAATGACATTAACCTTAATGAAAAAGTAGAGATTAATCAAGATGAC 1112
Db      ||| 410 AspAlaAspGlyLysIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 429
Qy      ||| 1113 ACCAAGCTATACATTAACCGTTATATAGGCAAGCGAACCGAAGAGAGATGCTACCTAC 1172
Db      ||| 430 ThrAsnAlaGlnIleThrValTyrMetGlyLysArgProGlnGlnGlnAlaSerTyr 449
Qy      ||| 1173 CATTTAGCT 1181
Db      ||| 450 HisLeuAla 452

```

```

; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
;
US-07-854-596B-28
;
Alignment Scores:
Pred. No.: 137e-189 Length: 499
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 1 Gaps: 0
;
US-09-940-235-9 (1-1541) x US-07-854-596B-28 (1-499)
Qy      ||| 33 ATAGCTGATCTCGAATATGCTACTAGATCGTCTTGTGAATAATACAGCCAAATGGTTGTT 92
Db      ||| 86 IleAlaGlyProGluThrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 105
Qy      ||| 93 AGCGTGTGCTGATCTGTTGAGGCGAGCAATCAAGACATTAAGCTTAAATTTTGAATC 152
Db      ||| 106 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 125
Qy      ||| 153 GATCTAACATCAAGACCTGCTCATGAGGAAAGACGAGAGAGGCTTAACTCAAAATCA 212
Db      ||| 126 AspLeuThrSerThrProAlaHisGlyGlyLysThrGlnGlnGlnLeuSerProLysSer 145
Qy      ||| 213 AAACATTTGCTACTAGATAGTGGCGGAGATGTCATTAACCTTGAGAAAGCTGACTTAATA 272
Db      ||| 146 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 165
Qy      ||| 273 AAGGCTATTCAGAGCAATTCGCTAAGCGCCAGATGAACAAGCACTTGTGAGTC 332
Db      ||| 166 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 185
Qy      ||| 333 ATTGATTTTGCAGCGATGCAACCATTCATGATCGAAAGCGCAAGGCTCTTGTGTCAC 392
Db      ||| 186 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 205
Qy      ||| 393 AAGATGATTCGCTAAGCTTGGCCGAGCCAACTGTCGAAGATTTTGTCAAGCGACAT 452
Db      ||| 206 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 225

```

RESULT 13
 US-07-854-596B-28
 ; Sequence 28, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M

```

QY 443 GTCCGCGTTAGACCATATTAAGAAAAACAATACAAACCAAGCAATCTGTGATGTG 512
Db 226 ValArgValArgProTyrIleGluysProIleGlnAsnGlnIleAlaYsSerValAspVal 245
QY 513 GAATTAATCTGTACAGTTACTCCCTTAAACCTCGATGACGATTTTCAGACCAAGCTTCAAA 572
Db 246 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGluLeuLys 265
QY 573 GATTAATAGCTATTGAAAAACATAGCTATCGGTGACACCACTGACATCTCAAGAAATTA 632
Db 266 AspThrIleValLeuLeuLysThrIleuAlaIleGlyAspThrIleThrSerGlnGluLeu 285
QY 633 GCTCAAGCAACAAAGCATTTTAAACAAAAAACCAACCAAGCTATACGATTTATGAACGTGAC 692
Db 286 AlaGlnAlaGlnSerIleLeuLeuLysThrIleAspProGluTyrThrIleTyrGlnIleAsp 305
QY 693 TCCTCAATGTCATCTATGACATGACATTTTCCGTAACGATTTTACCAATGATCAAGAG 752
Db 306 SerSerIleValIleThrIleAspAsnAspIlePheArgThrIleLeuProMetAspGlnIle 325
QY 753 TTTACTTACCGGTAAATAATGGGAAACAAGCTTATAGATCAATTAATAAATCTGGCTG 812
Db 326 PheIleTyrHisValIleValIleAsnArgGlnGlnAlaTyrGlnIleAsnLysSerGlyLeu 345
QY 813 AATGAAGAATTAACAAACACATGACCTGATCTCTGAGAAATATTAACGCTTAAAAAAGG 872
Db 346 AsnGluGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValIleuLysGly 365
QY 873 GAAAACCGTATGATCCCTTGAATGACGATCACTGAAACTGTCAACATCAAAATCGTT 932
Db 366 GluLysProTyrAspProPheAspArgSerHisIleuLysLeuPheThrIleLysTyrVal 385
QY 933 GATGTCGATACCAAGAAATTTGCTAAAAAGTGAGACCTTTACACTAGCAAGCTAAC 992
Db 386 AspValIleAsnThrAsnGlnLeuLeuLysSerGlnGlnIleuLeuThrIleAspGlnIleAsn 405
QY 993 TTAAGCTTCAGAGATTTATACGATCTCGATTAAGGCTTAACTACTCTACCAACAAATTC 1052
Db 406 LeuAspPheAspAspLeuTyrAspProAspArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 425
QY 1053 GATGCTTTTGTATTATGACATTAACCTTAACCTGAAAAGTAGAGATTAATCAGATGAC 1112
Db 426 AspAlaPheGlyIleMetAspTyrThrLeuThrGlnGlyValGlnAspAsnHisAspAsp 445
QY 1113 ACCAACCGTATATTAACGTTTATATGGCAAGCAACCCGAGAGAGAGAAATGCTAGCTAC 1172
Db 446 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyLysAsnAlaSerTyr 465
QY 1173 CATTAGCT 1181
Db 466 HisLeuAla 468

```

RESULT 14
US-07-854-596B-35

Sequence 35, Application US/07854596B

Patent No. 5434073

GENERAL INFORMATION:

APPLICANT: Dawson, Keith M

APPLICANT: Hunter, Michael G

APPLICANT: Czaplowski, Lloyd G

TITLE OF INVENTION: Proteins and nucleic acids

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dr. John J. McDonnell

STREET: Ten South Wacker Drive, Suite 3000

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854, 596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-35

Alignment Scores:
Pred. No.: 1,79e-189 Length: 859
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-9 (1-1541) x US-07-854-596B-35 (1-859)

QY 33 ATAGCTGCTGCTGAATGGCTACTAGATCGTCTTCTGTAATAATAACAGCAATTCGTGTT 92
Db 22 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 41
QY 93 AGCGTGTCTGTACTGTGAGGGGACGAATCAAGACATTAATGCTTAATTTTGAATC 152
Db 42 SerValAlaGlyThrValGlnGlnIleThrAsnGlnAspIleSerLeuLysPhePheGluIle 61
QY 153 GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 212
Db 62 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnIleuSerProLysSer 81
QY 213 AAACCATTTCTACTGATAGTGGCGGATGTACATATAAATTGAGAAAGCTGACTTACTA 272
Db 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 101
QY 273 AAGGCTATTCAAGAACAAATGATCGCTAAGTCCAGATGACGACTGTTAGAGTC 332
Db 102 LysAlaIleGlnIleGlnIleuLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 121
QY 333 ATTGATTTTTCAGACGATGCAACCATTAATGATGCAAGAGCAAGGCTACTTGTGTCAC 392
Db 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIleTyrPheAlaAsp 141
QY 393 AAAGATGTTGCTGTAACCTTGGCCGACCAACCTGTCCAGAAATTTTGTAGCGGACAT 452
Db 142 LysAspGlySerValThrLeuProThrGlnProValGlnIleuPheLeuLeuSerGlyHis 161
QY 453 GTCCGCGTTAGACCATTAAGAAAAACAATACAAACCAAGCAAGAAATGCTGATGTG 512
Db 162 ValArgValArgProTyrIleGlnLysProIleGlnAsnGlnIleLysSerValAspVal 181
QY 513 GAATTAATGTCATAGTTTACTCCCTTAAACCTGATGATGATTCGACCAAGGCTTCAAA 572
Db 182 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
QY 573 GATTAATAGCTATTGAAAAACATAGCTATCGGTGACACCACTGACATCTCAAGAAATTA 632
Db 202 AspThrIleLeuLeuLysThrIleuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 221
QY 633 GCTCAAGCAACAAAGCATTTTAAACAAAAAACCAACCAAGGCTATACGATTTATGAACGTGAC 692

```

```
Db 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 241
Qy TCCTCAATCGTCACATCATGACATATTCCTCGTACGATTTTACCAATGGATCAAGAG 752
Db 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 261
Qy TTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 812
Db 262 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 281
Qy 813 AATGAAGAAATAACAACACATGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 872
Db 282 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 301
Qy 873 GAAAGCCGATGATCCCTTTGATCGCAGTCATCTGAAACTGTTACCATCAATAATACGTT 932
Db 302 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
Qy 933 GATGTCGATACCAACGAATTGCTAAAGAGTGACGCTCTTAACAGCTAGCCGAACGTAAC 992
Db 322 AspValAsnThrAsnGluLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 341
Qy 993 TTAGACTTTCAGAGATTATACGATCCTCGTGATAGGCTTAAACTCTCTACCAATCTC 1052
Db 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 361
Qy 1053 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGAGGATATCAGATGAC 1112
Db 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 381
Qy 1113 ACCAACCTATCATACACGTTTATATGGCAGCGACCGACGAGGAGAGATGCTACTAC 1172
Db 382 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 401
Qy 1173 CATTTAGCT 1181
Db 402 HisLeuAla 404
```

RESULT 15

```
US-08-488-940-18
; Sequence 18, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 245 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-940-18

Alignment Scores:
Pred. No.: 2 53e-187 Length: 1194
Score: 1937.50 Matches: 390
Percent Similarity: 88.64% Conservative: 8
Best Local Similarity: 86.86% Mismatches: 45
Query Match: 71.34% Indels: 6
DB: 1 Gaps: 3

US-09-940-235-9 (1-1541) x US-08-488-940-18 (1-1194)
Qy 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCCTTCTGTAAATAACACCAATATGGTTGTT 92
Db 382 IleAlaGlyProGluTrpLeuLeuAspAlaProSerValAsnAsnSerGlnLeuVal 401
Qy 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 152
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuAlaPhePheGluIle 421
Qy 153 GATCTAACATCACGACCTGCTCATGGAGGAACACACAGCAAGCTTAAAGTCCAAATCA 212
Db 422 AspLeuThrSerAlaProAlaHisGlyGlyAlaThrGluGlnGlyLeuSerProAlaSer 441
Qy 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 272
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
Qy 273 AAGCTTATTCAGAAACAATTGATCGCTAAGCTCCACAGTAAACGACGACTCTTTGAGGTC 332
Db 462 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
Qy 333 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGCTCTACTTTGCTGAC 392
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
Qy 393 AAAGATGGTTCGGTAACTTGGCCGACCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 452
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 521
Qy 453 GTCCGCTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 512
Db 522 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
Qy 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACCAGGCTCTCAA 572
Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 560
Qy 573 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 632
Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 580
Qy 633 GCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGACCGTGAC 692
Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
Qy 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 752
Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
Qy 753 TTTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
Qy 813 AATGAAGAAATAACAACACATGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 872
Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660
```

```
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACTGTTCACCATCAATACGTT 932
Db |||||||
QY 661 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 680
Db |||||||
QY 933 GATGTCGATACCAACGAATTGCTAAAGAGAGCAGCTCTTAACAGCTAGCGACGTAAC 992
Db |||||||
QY 681 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
Db |||||||
QY 993 TTAGACTTCAGAGATTATACGATCTCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1052
Db |||||||
QY 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
Db |||||||
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAAAAGTAGAGGATAATCACGATGAC 1112
Db |||||||
QY 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
Db |||||||
QY 1113 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGATGCTAGCTAC 1172
Db |||||||
QY 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
Db |||||||
QY 1173 CATTTAGCTGGTGGTGGCCAGCGCAACAGATTGTACCCATAGCTGAGAAGTGTTTTGAT 1232
Db |||||||
QY 761 HisLeuAla-----TyrAspAlaAspArgTyrThrGluGluArgGluValTyrSer 778
Db |||
QY 1233 CATGCTGCGGGACTTCCTATGCTGCGGAGAAACGTTGGAGAGAGCCCTACCAAGGCTGG 1292
Db |||
QY 779 TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLysAsnAsnSerGln 798
Db |||
QY 1293 ATCATGCTAGATTGT-----ACTTGCCTGGGAGAGGACGCGACGCATCATTGCACT 1346
Db |||
QY 799 LeuValValSerValAlaGlyThrValGluGlyThrAsnGln-AspIleSerLeuLysPhe 818
Db |||
QY 1347 TCTAGAAATAGATGCAACGATCAGG 1371
Db |||
QY 818 ePheGluIleAspLeuThrSerArg 826
Db |||
```

Search completed: January 28, 2006, 02:40:58
Job time : 40.9832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:51:16 ; Search time 100.774 Seconds
(without alignments)
12778.538 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 2716
Sequence: 1 ttgtttaacttaagaagg.....ccttcaccgatgttcgtag 1541

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US0940235/runat_27012006_144219_27635/app_query.fasta_1.7708
-DB=Published Applications AA Main -OFMT=fasta -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -USER=US0940235 @CGN 1 1 805 @runat_27012006_144219_27635
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA_Main.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1984	73.0	413	4	US-10-360-101-264
2	1984	73.0	413	3	Sequence 264, App
3	1963	72.3	415	5	US-09-940-235-2
4	1958	72.1	414	4	US-10-988-943-1
5	1947	71.7	414	4	US-10-300-215-252
6	1928	71.0	413	3	US-10-300-215-253
7	1794	66.1	440	5	US-09-919-703-12
8	632	23.3	259	3	US-10-474-792-658
9	632	23.3	463	4	US-09-940-235-4
10	632	23.3	463	5	US-10-144-194A-52
11	632	23.3	642	4	US-10-491-566-52
12	632	23.3	642	5	US-10-741-601-354
					Sequence 1066, App

13	632	23.3	657	4	US-10-741-601-359
14	632	23.3	657	5	US-10-741-600-1072
15	632	23.3	984	4	US-10-741-601-356
16	632	23.3	984	5	US-10-741-600-1069
17	632	23.3	2220	4	US-10-236-392-4
18	632	23.3	2296	4	US-10-741-601-363
19	632	23.3	2296	5	US-10-741-600-1075
20	632	23.3	2320	4	US-10-279-733-8
21	632	23.3	2320	4	US-10-236-392-2
22	632	23.3	2328	4	US-10-171-311-64
23	632	23.3	2328	4	US-10-236-031B-70
24	632	23.3	2328	4	US-10-374-979-98
25	632	23.3	2328	4	US-10-182-936A-98
26	632	23.3	2328	5	US-10-477-238A-677
27	632	23.3	2328	5	US-10-680-287A-677
28	632	23.3	2328	5	US-10-477-173-677
29	632	23.3	2355	4	US-10-144-194A-104
30	632	23.3	2355	4	US-10-360-101-235
31	632	23.3	2355	4	US-10-447-161-3
32	632	23.3	2355	4	US-10-734-564-94
33	632	23.3	2355	4	US-10-741-601-357
34	632	23.3	2355	4	US-10-741-601-366
35	632	23.3	2355	5	US-10-451-566-104
36	632	23.3	2355	5	US-10-741-600-1067
37	632	23.3	2355	5	US-10-741-600-1078
38	632	23.3	2355	5	US-10-852-335A-147
39	632	23.3	2355	5	US-10-287-436A-436
40	632	23.3	2355	5	US-10-287-436A-1137
41	632	23.3	2355	6	US-11-040-130-28
42	632	23.3	2386	3	US-09-961-403-1
43	632	23.3	2386	4	US-10-741-601-360
44	632	23.3	2386	4	US-10-788-792-206
45	632	23.3	2386	5	US-10-618-281-32

ALIGNMENTS

RESULT 1
US-10-360-101-264
; Sequence 264, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:			
Pred. No.:	8.21e-177	Length:	413
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.05%	Indels:	0
DB:	4	Gaps:	0

US-09-940-235-9 (1-1541) x US-10-360-101-264 (1-413)

QY 33 ATAGCTGGTCTCAATGGCTACTAGATCGTCCTTCTTAATAACACCAATGGTGTGT 92
Db 1 ILEALGlyProGluTrpLeuLeuAspArgProSerValAsnSerGlnLeuVal 20


```

QY 453 GTGCGGTAGACCATATAAGAAAACCAATACAAACCAAGGAAATCTGTTGATGTG 512
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAACAGGCTCTCAA 572
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATACCTAAGCTATTGAAAACACTAGCTATCCGTGACACCACTCATCTCAAGAAATTA 632
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 752
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTTACTTACCGTGTAAATAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 813 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 872
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTGAACTGTTCCACCATCAATAGCTT 932
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGTCGATACCAACCAATTTGCTTAAAGAGTGACAGCTCTTAACAGCTAGCAACGTAAC 992
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGAGATTATACGATCTCGTGATAAGCTAAACTACTCTACACAAATCTC 1052
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATCAAGTAC 1112
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAAGCGTATCATACCGTTTATATGGCAAGCGCCGAGAGAGATGCTAGCTAC 1172
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
Db 381 HisLeuAla 383

```

RESULT 3

```

US-10-988-943-1
; Sequence 1, Application US/10988943
; Publication No. US20050176085A1
; GENERAL INFORMATION:
; APPLICANT: Center for Genetic Engineering and Biotechnology
; TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
; TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
; TITLE OF INVENTION: COMPLEX MIXTURES.
; FILE REFERENCE: Proteomics CU2003-269
; CURRENT APPLICATION NUMBER: US/10/988,943
; CURRENT FILING DATE: 2004-11-15
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-988-943-1
Alignment Scores:

```

```

Pred. No.: 7,66e-175 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 72.28% Indels: 0
DB: 5 Gaps: 0

```

US-09-940-235-9 (1-1541) x US-10-988-943-1 (1-415)

```

QY 30 ATGATAGTCTGCTCTGCTGCTAGATCGTCTCTCTGCTAAATAACAGCAATGCTT 89
Db 1 MetIleAlaGlyProGluIrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 90 GTTAGCCGTGCTGCTACTGTTGAGCGGACGAATCAAGACATTAGTCTTAAATTTTTGAA 149
Db 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGlu 40
QY 150 ATCGATCTAATCAATCACGACCTGCTCATGAGGAAGACAGACGAGGCTTAAGTCCAAA 209
Db 41 IleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLys 60
QY 210 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATCTGAGAAAGCTGACTTA 269
Db 61 SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu 80
QY 270 CTAAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAACGACGACTCTTTGAG 329
Db 81 LeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlu 100
QY 330 GTCAATTGATTTGCAAGCGATGCAACCAATTACTGATGAAACGCAAGGCTACTTTGCT 389
Db 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAla 120
QY 390 GACAAAGATGGTTCGGTAACCTTCCCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGGA 449
Db 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly 140
QY 450 CATGTGCGCTAGACCATATAAGAAAACCAATACAAACCAAGCGCAATCTGTTGAT 509
Db 141 HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 160
QY 510 GTGGAATATATCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGCTC 569
Db 161 ValGluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeu 180
QY 570 AAGATATCTAAGCTATTGAAAACACTAGCTATCGGTGACCACTCATCTCTCAAGAAATTA 629
Db 181 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 630 CTAGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 689
Db 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArg 220
QY 690 GACTCTCTCAATCTGCTCATGACAAATGACATTTTCGTAGCTAGCTTTTACCAATGATCAA 749
Db 221 AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240
QY 750 GAGTTTACTTACCGTGTATAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGCT 809
Db 241 GluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGly 260
QY 810 CTGAATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAA 869
Db 261 LeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLys 280
QY 870 GGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTAACCTCAACCAATAC 929
Db 281 GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 300
QY 930 GTTGATGTCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAAACAGCTAGCAACGT 989
Db 301 ValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320

```

QY 990 AACTTAGCTTCAGAGATTATACAGTCTCGGTGATAGGCTAAACTACTCTTACAACAT 1049
|||||
Db 321 AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 340
QY 1050 CTCGATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAGTAATACAGAT 1109
|||||
Db 341 LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 360
QY 1110 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGATGCTAGC 1169
|||||
Db 361 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 380
QY 1170 TACCATTTAGCT 1181
|||||
Db 381 TyrHisLeuAla 384

RESULT 4

US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252

Alignment Scores:
Pred. No.: 2,25e-174 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 72.09% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-9 (1-1541) x US-10-300-215-252 (1-414)

QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTTCTGTAATAACAGCAATTTGGTTGT 92
|||||
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 93 AGCGTTGCTGTTACTGTTGAGGGGAGCATCAAGCATAGTCTTAATTTTGAATC 152
|||||
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 153 GATCTAAACATCACGACCTCTCATGGAGGAAACAGACAGAGCGCTTAAGTCCAAAATCA 212
|||||
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60

QY 213 AAACCAATTTGCTACTGATGTCGGCGATGTCACATAAACTTGAGAAAGCTGACTTACTA 272
|||||
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 273 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCCACAGTAACGACGACTACTTTGAGTC 332
|||||
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 333 ATTGATTTTCCAGCCGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 392
|||||
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAAGATGGTTTCGGTAAACCTTTGCGACCCCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT 452
|||||
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAACCAACGAAATCTGTTGATGTG 512
|||||
Db 141 ValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATCTGTACAGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
|||||
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTA 632
|||||
Db 181 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 633 GCTCAAGCACAAAGCATTTTAAACAAAAACCCAGGCTATACGATTTATGAAACGTCAC 692
|||||
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATCGTCTCATGACAATGACATTTTCGTCAGATTTTTCACATGAGATCAAGAG 752
|||||
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 753 TTTACTTTACCGTCTTAAATAATCGGAACAGCTTATAGGATCAATAAATAATCTGGCTG 812
|||||
Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
QY 813 AATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
|||||
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCCGTCATCATCCCTTTGATCGCAGTCACCTTGAAACTGTTTCACCATCAATACGTT 932
|||||
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGACAGCTCTTAACAGCTAGGACGCTAAC 992
|||||
Db 301 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGAGATTTTATACAGATCTCTCGGTGATAAGGCTAAACTACTCTCAACAATCTC 1052
|||||
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAGGATAATACGATGAC 1112
|||||
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACGGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGATGCTAGTAC 1172
|||||
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
|||||
Db 381 HisLeuAla 383

RESULT 5

US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1

GENERAL INFORMATION:
 ; APPLICANT: CARR, Francis Joseph
 ; APPLICANT: ADAIR, Fiona Suzanne
 ; APPLICANT: HAMILTON, Anita Anne
 ; APPLICANT: CARTER, Graham
 ; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
 ; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
 ; FILE REFERENCE: MER-104-CON.1
 ; CURRENT APPLICATION NUMBER: US/10/300,215
 ; CURRENT FILING DATE: 2002-11-20
 ; PRIOR APPLICATION NUMBER: US 09/438,136
 ; PRIOR FILING DATE: 1999-11-10
 ; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
 ; PRIOR FILING DATE: 1998-05-21
 ; PRIOR APPLICATION NUMBER: GB 9710480.6
 ; PRIOR FILING DATE: 1997-05-21
 ; PRIOR APPLICATION NUMBER: GB 9716197.0
 ; PRIOR FILING DATE: 1997-07-31
 ; PRIOR APPLICATION NUMBER: GB 9725270.4
 ; PRIOR FILING DATE: 1997-11-28
 ; PRIOR APPLICATION NUMBER: GB 9807751.4
 ; PRIOR FILING DATE: 1998-04-14
 ; PRIOR APPLICATION NUMBER: US 60/067,235
 ; PRIOR FILING DATE: 1997-12-02
 ; NUMBER OF SEQ ID NOS: 254
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 253
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Modified strep protein
 US-10-300-215-253

Alignment Scores:
 Pred. No.: 2,42e-173 Length: 414
 Score: 1947.00 Matches: 376
 Percent Similarity: 98.69% Conservative: 2
 Best Local Similarity: 98.17% Mismatches: 5
 Query Match: 71.69% Indels: 0
 DB: 4 Gaps: 0

US-09-940-235-9 (1-1541) x US-10-300-215-253 (1-414)

QY 33 ATAGTGTGCTGCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATGCTGTT 92
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 93 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 153 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGCAAGCTTAAGTCCAAATCA 212
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 213 AAACATTGCTACTGATAGTGGCGGATGTCATAAATCGAGAAAGCTGACTACTA 272
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
 QY 273 AAGCGTATTCAAGCAATGTAGTCGTAAAGTCCAGTAAACGACACTACTTTGAGTTC 332
 Db 81 LysAlaLysGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 333 ATTGATTTTGAAGCGATGCAACCACTTACTGATCGAAACGGCAAGCTTACTTTGCTGAC 392
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 393 AAAGATGGTTCGGTAACCTTGCGGACCCCACTGTCCTCAAGAATTTTGTCTAAGCGGACAT 452
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 453 GTGCGGGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCAAAATCTGTTGATGTC 512

Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 513 GAATATATCTGTACAGATTCTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 572
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheAspArgProGlyLeuLys 180
 QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCCGTGACACCATCACATCTTCAAGAATTACTA 632
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
 QY 633 GCTCAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTAGTACGATTTTACCAATGATCAAGAG 752
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 753 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
 QY 813 AATGAAGAAATAAACAACACTGACCTCATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTTTCACCATCAAAATACGTT 932
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysPheVal 300
 QY 933 GATGTCGATACCAACGAATTCGTAAGAGTGAGCAGCTCTTAACAGCTAGCGACGTAAC 992
 Db 301 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 993 TTAGACTTCAGAGATTTTATACGATCTCTCGTGAAGCTAACTACTCTACAACAATCTC 1052
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1053 GATGCTTTGCTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACAGATGAC 1112
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1113 ACCAACCGTATCAACCGTTTATATGCGCAAGCGACCGCAAGAGAGAGATGCTAGCTAC 1172
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1173 CATTTAGCT 1181
 Db 381 HisLeuAla 383
 RESULT 6
 ; US-09-919-703-12
 ; Sequence 12, Application US/09919703
 ; Patent No. US20020165129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krystal, Gerald
 ; APPLICANT: Rabkin, Simon W.
 ; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
 ; TITLE OF INVENTION: Cell Death
 ; FILE REFERENCE: 50216/003004
 ; CURRENT APPLICATION NUMBER: US/09/919,703
 ; CURRENT FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 09/294,457
 ; PRIOR FILING DATE: 1999-04-19
 ; PRIOR APPLICATION NUMBER: US 08/759,599
 ; PRIOR FILING DATE: 1996-12-05
 ; PRIOR APPLICATION NUMBER: US 60/008,233
 ; PRIOR FILING DATE: 1995-12-06
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 413

db	280	GlyGluysProTyAspProPheAspArgSerHisLeuIysLeuPheThrIleLysTyr	299
Qy	930	GTTGATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGT	989
Db	300	ValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg	319
Qy	990	AACTTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTACTCTACACAAT	1049
Db	320	AsnLeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsn	339
Qy	1050	CTCGATGCTTTGTGTTATATGGAATACTTAACTGGAAGAGTAGAGGATAATCAGAT	1109
Db	340	LeuAspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluLysAspAsnHisAsp	359
Qy	1110	GACACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGC	1169
Db	360	AspThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSer	379
Qy	1170	TACCAT	1175
Db	380	TyrHis	381

US-10-474-792-658
; Sequence 658, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 658
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-658

Alignment Scores:		
Pred. No.:	5.5e-159	Length:
Score:	1794.00	Matches:
Percent Similarity:	83.99%	Conservative:
Best Local Similarity:	99.03%	Mismatches:
Query Match:	66.05%	Indels:
DB:	5	Gaps:
		0
		440

US-09-940-235-9 (1-1541) X US-10-474-792-658 (1-440)

Qy	33	ATAGCTGGTCTCGAATCGCTACTAGATCGTCTCTCTGTAATAACAGCCAAATTGGTTGT	92
Db	27	IleAlaGLtyrGIyTrpLeuProAspArgProproIleAsnAsnSerGlnLeuVal	46
Qy	93	AGCGTTCGTGCTACTGTTTGAGGGGCAAGATCAAGACATTAGCTCTTAAATTTTTTCGAATC	152
Db	47	SerMetaLaGLylleValGLUGlyThrAspLysLysValPheIleasnPheGluIle	66
Qy	153	GATCTTAACATCCAGCACTGCTCATGGAGAAAGACAGACGCAAGCGCTTAAAGTCCAAAATCA	212
Db	67	AspLeuThrSerGlnProAlahiesGLyLysThrGIuGInGLyLeuSerProLysSer	86
Qy	213	AAACCATTTCCTACTGATAGTGGCGCGATGCACATAAATTTGAGAAAGCTCACTTACTA	272
Db	87	LysProPheAlaThrAspAsnGlyalaMetProHisLysLeuGLyLysAlaaspLeuLeu	106
Qy	273	AAGGCTATTCCAAGAACAAATTGATCGCTAACGTCCTCACAGTAACGACGACTACTTTGAGGTC	332
Db	107	LysAlalileGlnLysLeuIleAlaAsnValHisSerAsnAspGLYtyrPheGluVal	126

33	ATAGCTGGTCTCGAATGGCTACTAGATCGTCTCTCTGTGTAATAACACGACCAATGGTGTGTT	99
1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
93	AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC	152
21	SerValAlaGlyThrVal---GlyThrAsnGlnAspIleSerLeuLysPheGluIle	39
153	GATCTAACATCACGACCTGCTCATGGAGAAAGACAGACAAAGGCTTAAGTCCAAAATCA	212
40	AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	59
213	AAACGATTTGCTACTGATAGTGGCGCGATGCACATAACTTCAGAGAACTGACTTACTTA	272
60	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu	79
273	AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAGACGACTACTTTGAGGTC	332
80	LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	99
333	ATTGATTTTGCACGGATGCAACCAATTACTGATTCGAACACGGCAGGCTCTACTTTGCTGAC	392
100	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	119
393	AAACATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAAAGATTTTGTCTAAGCGGACAT	452
120	LysaspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	139
453	GTGCGCGTTAGACCATATAAGAAAAAACCAATACAAAAACCAAGCGAAATCTGTGTGATGTG	512
140	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	159
513	GAATATACTGTACAGTTTACTCCCTTAACCTGTAGCAGTATTCAGACCAGGCTCTCAA	572
160	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	179
573	GATCTAAGCTATTGAAAACACTAGCTATCGGTGCACCCATCACATCATCTCAAGAATTACTTA	632
180	LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	199
633	GCTCAAGCACAAAGCAATTTTAAACAAAACCAACCGAGCTATACGATTTATGAACGTGAC	692
200	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	219
693	TCCCTCAATCGTCACCTCATGCACATGCATTTTCCGTACGATTTTACCAATGGATCAAGAG	752
220	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	239
753	TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	812
240	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	259
813	AATCAAGAAATAACCAACACTGCACCTGATCTCTGAGAATAT---TACGTCCTTAAAAA	869
260	AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluLysTyrValLeuLysLys	279
870	GGGGAAAGCCGATGATGCCCTTTGATCGCAGTCATCTGAAACTGTTTACCACATCAATAC	929

```
QY 333 ATTGATTTTCAGCGGATGCAACCATTAATGATGAAACGCAAGGCTACTTTGCTGAC 392
Db 127 ILeAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAGATGCTTCGTTAACTTCGCGACCAACCTGTCACAGATTTTTCGCTTAAGCGGACAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 166
QY 453 GTCGGCTTAGACCATATAAGAAAAACCAATACAAACCAACCAAGCAATCTGTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCAAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 633 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGCTATACGATTTATGACGTCAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGCAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTACCGTGTAAATTCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY 813 AATGAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrValLeuLysGlnGly 306
QY 873 GAAACCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCACATCAATAGT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCGATACCAACGAATTTGCTAAAGATGACAGCTCTTAAACAGCTAGCGAAGCTAAC 992
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGCTTCAGAGATTTATAGATCTCTCGTGATAAGCTAAACTACTCTCAACAACATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1112
Db 367 AspAlaPheAspLysMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCCAAGAGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
```

RESULT 8

```
US-09-940-235-4
; Sequence 4, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
```

```
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-235-4
Alignment Scores:
Pred. No.: 4,33e-50 Length: 259
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 3 Gaps: 3
US-09-940-235-9 (1-1541) x US-09-940-235-4 (1-259)
QY 1143 AAGCGACCCGAGGAGAGAAATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 124 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 142
QY 1191 CAGCGGCAA---CAGATTGTATCCCATAGCTGAGAAAGTGTGTTGATCATGCTGGGACT 1247
Db 143 LysGlyGluTyrThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 162
QY 1248 TCCTATCTGTCGGGAGAAAGCTGGAGAACCCCTACCAAGCTGGATGATGTTAGATTGT 1307
Db 163 SerTyrValValGlyGluThrTyrPglLysProTyrGlnGlyTyrMetMetValAspCys 182
QY 1308 ACTTGCCTGGGAGAGGAGCGGACGATCATCTTGCACCTCTAGAAATAGATGCAACGAT 1367
Db 183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
QY 1368 CAGGACCAAGGACATCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTyrSerLysLysAspAsnArgGly 222
QY 1428 AACCTGCTCCATGTCATCTGCACAGGCAACGGCCGAGAGAGTGGAAAGTGTGAGAGGCAC 1487
Db 223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTyrLysCysGluArgHis 242
QY 1488 ACCTCTCTGAGACACCATCGAGCGGATCTGGCCCTTTCACCGATGTTCTGT 1538
Db 243 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 259
RESULT 9
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
```



```
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGTCAGACCATCGCGGATCTGGCCCTTACCGATGTTGCT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 12
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Alignment Scores:
Pred. No.: 5,77e-50 Length: 642
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 5 Gaps: 3

US-09-940-235-9 (1-1541) x US-10-741-600-1066 (1-642)
QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCAATTTA-----GCTGTGTGGTGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGCGCAA--CAGATTGTACCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACT 1247
Db 174 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaGlyThr 193
QY 1248 TCCTATGTGTCGGAGAACGTCGGAGAACCCCTACCAAGCTCGATGATGTTAGTGT 1307
Db 194 SerTyValValGlyGluThrTrpGluLysProTyArgGlyGlnGlyTrpMetValAspCys 213
QY 1308 ACTTGCCTGGGAGAGCGGACGCATCTGCACTTCTAGAAATAGATGCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGACATCCTATAGAAATGGAGACACCTGGAGACAGGAGGATATCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCCAGTGCATCTATAGAAATGGAGACACCGCGGAGAGTGGAAAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGTCAGACCATCGCGGATCTGGCCCTTACCGATGTTGCT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 13
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
```

```
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Alignment Scores:
Pred. No.: 5,81e-50 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 4 Gaps: 3

US-09-940-235-9 (1-1541) x US-10-741-601-359 (1-657)
QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCAATTTA-----GCTGTGTGGTGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGCGCAA--CAGATTGTACCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACT 1247
Db 174 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaGlyThr 193
QY 1248 TCCTATGTGTCGGAGAACGTCGGAGAACCCCTACCAAGCTCGATGATGTTAGTGT 1307
Db 194 SerTyValValGlyGluThrTrpGluLysProTyArgGlyGlnGlyTrpMetValAspCys 213
QY 1308 ACTTGCCTGGGAGAGCGGACGCATCTGCACTTCTAGAAATAGATGCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAGGACATCCTATAGAAATGGAGACACCTGGAGACAGGAGGATATCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCCAGTGCATCTATAGAAATGGAGACACCGCGGAGAGTGGAAAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGTCAGACCATCGCGGATCTGGCCCTTACCGATGTTGCT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 14
US-10-741-600-1072
; Sequence 1072, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1072
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1072

Alignment Scores:
Pred. No.: 5,81e-50 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
```


DB:	5	Gaps:	3
US-09-940-235-9 (1-1541) x US-10-741-600-1072 (1-657)			
QY	1143 AAGCGACCCGAAGGAGAGAATGCTAGCTACCATTTA-----GCTGGTGGTGGC	1190	
	:::		
Db	155 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly	173	
	:::		
QY	1191 CAGCGCGCAA---CAGATTGTACCCATAGCTGAGAAAGTGTTTTGATCATGCTGCTGGGACT	1247	
	::: :::		
Db	174 LysGlyGluThrThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr	193	
	:::		
QY	1248 TCCTATCTGTCGCGAGAAACGTGGGAAAGCCCTACCAAGGCTGGAGTGTGTAGATTGT	1307	
	:::		
Db	194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTyrMetMetValAspCys	213	
	:::		
QY	1308 ACTTGCTCGGAGAGCGAGCGACGCATCACATTGCATCTTCTAGAAATAGATGCAACGAT	1367	
	:::		
Db	214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp	233	
	:::		
QY	1368 CAGGACACAGGACATCCTATAGATTGGGAGACACCTGGAGCAGAAGGATATTCGAGGA	1427	
	:::		
Db	234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerIlyAspAsnArgGly	253	
	:::		
QY	1428 AACCTGCTCAGATGTCATCTGCACAGGCCAACCGCCGAGCAGTGGAGTGTGAGAGGCAC	1487	
	:::		
Db	254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluThrIlyCysGluArgHis	273	
	:::		
QY	1488 ACTTCTGTGAGACCAACATCAGCGGATCTGGCCCTTCACCGATGTTGCT	1538	
	:::		
Db	274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg	290	
	:::		

RESULT 15

```

US-10-741-601-356
; Sequence 356, Application US/10741601
; Publication No.: US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 356
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-356

```

Alignment Scores:

Pred. No.:	6,6e-50	Length:	984
Score:	632.00	Matches:	117
Percent Similarity:	87.59%	Conservative:	3
Best Local Similarity:	85.40%	Mismatches:	11
Query Match:	23.27%	Indels:	6
DB:	4	Gaps:	3

US-09-940-235-9 (1-1541) x US-10-741-601-356 (1-984)

Qy	1143	AAGCGACCCGAAGAGAGAAATGCTAGCTACCATTTA-----GCTGGTGTGGC	11190
		:::	
Db	155	ArgArgProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly	173
Qy	1191	CAGCGCAA---CAGATGTACCATAGCTGAGAGTGTTCATCATGCTGTGGGACT	1247
		::: :::	
Db	174	LysGlyGluIrrpThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr	193
Qy	1248	TCCTATGTGTGCGGAGAAACGTGGGAGAACCCCTACCAAGCGTGGATGATGGTAGATTGT	1307
Db	194	SerTyrrValValGlyGluThrTrpGluLysProTyrrGlnGlyTrpMetMetValAspCys	213

Qy	1308	ACTTGCCTGGGAGAGGACGGGACGGACGATC	ACTTGCCTCTAGAATAAGATGCAACGAT	1367
Db	214	ThrCysLeuGlyGluGlySerGlyArgIle	ThrCysThrSerArgAsnArgCysAsnAsp	233
Qy	1368	CAGGACACAAAGGACATCCTATAGAAATGGAGACACCTTGGAGCAAGAGGAGGATATCCAGGA	1427	
Db	234	GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerIlysAspAsnArgGly	253	
Qy	1428	AACCTGCTCCAGTGATCTGCACAGGCAACGGCCGAGGACAGTGGAGTCTGAGAGGCAC	1487	
Db	254	AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpIysCysGluArgHis	273	
Qy	1488	ACTTCTGTGACAGCACCATCATCAGCGGATCTGGGCCCTTCACCGATGTTGCT	1538	
Db	274	ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg	290	

Search completed: January 28, 2006, 02:56:30
Job time : 113.774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 02:09:11 ; Search time 4.46194 Seconds
(without alignments)
7479.985 Million cell updates/sec

Title: US-09-940-235-9

Perfect score: 2716

Sequence: 1 ttgttgtaacttaagaagg.....ccttcacgatgttcgttag 1541

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 151242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09940235/runat_27012006_144220_27676/app.query.fasta_1.7708
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=ext -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09940235.ecgn_1_1 @runat_27012006_144220_27676
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:

1: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pdb:
2: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pdb:
3: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pdb:
4: /cgn2_6/prodata/2/pubppaa/PCT_NEW_PUB.pdb:
5: /cgn2_6/prodata/2/pubppaa/US09_NEW_PUB.pdb:
6: /cgn2_6/prodata/2/pubppaa/US10_NEW_PUB.pdb:
7: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pdb:
8: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pdb:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1984	73.0	414	6	US-10-631-558-2
2	1928	71.0	413	7	US-11-032-951-12
3	632	23.3	259	6	US-10-631-558-4
4	632	23.3	300	7	US-11-006-119-31
5	632	23.3	642	6	US-10-995-561-631
6	632	23.3	657	6	US-10-995-561-622
7	632	23.3	984	6	US-10-995-561-629
8	632	23.3	2296	6	US-10-995-561-633
9	632	23.3	2355	6	US-10-995-561-623
10	632	23.3	2355	6	US-10-995-561-627

11	632	23.3	2384	6	US-10-821-234-1545
12	632	23.3	2386	6	US-10-995-561-626
13	265.5	9.8	1341	6	US-10-995-561-621
14	262.5	9.7	693	6	US-10-995-561-632
15	262.5	9.7	1315	6	US-10-995-561-630
16	262.5	9.7	1348	6	US-10-995-561-624
17	247	9.1	1259	6	US-10-995-561-625
18	208	7.7	1286	6	US-10-995-561-628
19	189.5	7.0	847	6	US-10-995-561-634
20	120	4.4	1992	7	US-11-013-759-3
21	120	4.4	1992	7	US-11-013-759-13
22	120	4.4	2047	7	US-11-013-759-4
23	120	4.4	2047	7	US-11-013-759-7
24	116.5	4.3	1647	7	US-11-052-554A-260
25	115	4.2	21	7	US-11-032-951-7
26	113	4.2	470	6	US-10-485-517-319
27	113	4.2	801	7	US-11-174-150-29
28	113	4.2	801	7	US-11-124-368A-292
29	112	4.1	408	7	US-11-016-564-5
30	109.5	4.0	1417	7	US-11-052-554A-8
31	109.5	4.0	1448	6	US-10-485-517-212
32	109	4.0	2340	7	US-11-052-554A-171
33	108.5	4.0	463	6	US-10-793-626-960
34	108.5	4.0	2314	7	US-11-013-759-11
35	106	3.9	876	7	US-11-077-550-82
36	106	3.9	876	7	US-11-077-550-106
37	106	3.9	876	7	US-11-077-550-108
38	106	3.9	2710	7	US-11-051-453-41
39	105	3.9	824	6	US-10-821-234-1008
40	105	3.9	876	7	US-11-077-550-128
41	105	3.9	877	7	US-11-077-550-126
42	105	3.9	877	7	US-11-077-550-130
43	105	3.9	881	7	US-11-077-550-124
44	105	3.9	902	7	US-11-077-550-132
45	105	3.9	912	7	US-11-077-550-116

ALIGNMENTS

RESULT 1
US-10-631-558-2
; Sequence 2, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-631-558-2

Alignment Scores:


```
Db 80 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 99
QY 333 ATTGATTTTCAAGCGATGCAACCATCTACTGATCGAAACGCGAGGCTCTCTTCTGCTAC 392
Db 100 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIleValTyrPheAlaAsp 119
QY 393 AAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
Db 120 LysAspGlySerValThrLeuProThrGlnProValGlnProValGlnPheLeuLeuSerGlyHis 139
QY 453 GTGCGGTAGACCATATAAGAAAAACCAATACAAAAACCAAGGAAATCTGTTGATGTG 512
Db 140 ValArgValArgProTyrIleGluLeuProIleGlnAsnGlnAlaLysSerValAspVal 159
QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 572
Db 160 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
QY 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 632
Db 180 LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 199
QY 633 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
QY 693 TCTTCAATCGTCACTCATGCAATGATCATTTTCGTACGATTTTACCAATGGATCAAGAG 752
Db 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
QY 753 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db 240 PheThrTyrArgValLysAsnAsgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 259
QY 813 ANTGAAGAAATAAACAACTGACCTGATCTCTGAGAAATAT---TAGCTCTTTAAAAAA 869
Db 260 AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluTyrLysTyrValLeuLysLys 279
QY 870 GGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTTTACAGCTAGCGACGT 929
Db 280 GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 299
QY 930 GTTGATGTGATACCAACCAAGTGTCTAAAGTGAGCAGCTCTTAACAGCTAGCGACGT 989
Db 300 ValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 319
QY 990 AACTTAGACTTCAGAGATTTATACGATCCTCTGATAAAGGCTAAACTTACTCTACAACAAT 1049
Db 320 AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 339
QY 1050 CTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGATAGAGGATAATCAGAT 1109
Db 340 LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 359
QY 1110 GACCAACCGGTATCATACCGTTTATATGGCAAGCGACCCAGCAGAGAGATGCTAGC 1169
Db 360 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 379
QY 1170 TACCAT 1175
Db 380 TyrHis 381
```

RESULT 3

```
US-10-631-558-4
; Sequence 4, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
```

RESULT 4

```
US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ndao, Momar
; APPLICANT: Ward, Brian
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: Perichon, Regis
```

```
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-558-4

Alignment Scores:
Pred. No.: 2,76e-40 Length: 259
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

US-09-940-235-9 (1-1541) x US-10-631-558-4 (1-259)
QY 1143 AAGCGACCCGAGGAGAGAGATGTAGCTACCATTTA-----GCTGGTGGTGGC 1190
Db 124 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 142
QY 1191 CAGCGCAA---CAGATGTGTACCATAGCTGAGAGTGTGTTTGTATCATGCTGGGACT 1247
Db 143 LysGlyGluTyrThrCysLysProIleAlaGluLysCysPheAspHisAlaGlyThr 162
QY 1248 TCCTATGTGTGCGGAGAAACCTGGAGAACCTACCAAGGCTGATCATGTGATCATGTGT 1307
Db 163 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetValAspCys 182
QY 1308 ACTTGCCTGGAGAGGAGCGACGATCATCTTGCATCTTAGAAAATAGATGCAACGAT 1367
Db 183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
QY 1368 CAGGACACAGGACATCTATAGAAATTGGAGACACCTGGAGACAGAGGATAATCGAGGA 1427
Db 203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 222
QY 1428 AACCTGCTCCAGTGCATCTGCACAGGCAACCGCGAGGAGAGTGTGAGTGTGAGAGGCAC 1487
Db 223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
QY 1488 ACCTCTGTGACACACATCATCGAGCGGATCTGGCCCTTCCACCGATGTTCTGT 1538
Db 243 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 259
```

APPLICANT: CIPHERGEN BIOSYSTEMS, INC.
TITLE OF INVENTION: Serum Biomarkers for Chagas Disease
FILE REFERENCE: 016866-012130US
CURRENT APPLICATION NUMBER: US/11/006,119
CURRENT FILING DATE: 2004-12-06
PRIOR APPLICATION NUMBER: US 60/527,153
PRIOR FILING DATE: 2003-12-05
PRIOR APPLICATION NUMBER: US 60/565,093
PRIOR FILING DATE: 2004-04-22
PRIOR APPLICATION NUMBER: US 60/625,519
PRIOR FILING DATE: 2004-11-06
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 31
LENGTH: 300
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
US-11-006-119-31

Alignment Scores:
Pred. No.: 2,81e-40 Length: 300
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

US-09-940-235-9 (1-1541) x US-11-006-119-31 (1-300)

```
QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----CCTGGTGTGGC 1190
Db 124 ArgArgProHis---GluThrGlyGlyrMetLeuGluCysValCysLeuGlyAsnGly 142
QY 1191 CAGGCGCAA---CAGATTGTACCATAGCTGAGAGTGTGTTGATCATCTGCTGGGACT 1247
Db 143 LysGlyGluThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 162
QY 1248 TCCTATGTGTGGGAGAAACGTGGGAGAACCCCTACCAAGCTGGAGTGGTAGATTGT 1307
Db 163 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 182
QY 1308 ACTTGCTGGGAGAGCGGACGCATCATTGCATCTTAGAAATAGATGCAACGAT 1367
Db 183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
QY 1368 CAGACACAAAGACATCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 222
QY 1428 AACCTGTCCAGTGCATCTGCACAGCAACCGCCGAGGAGAGTGGAGTGTGAGAGGCAC 1487
Db 223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
QY 1488 ACCTCTGTGCAGACACATCGAGCGGATCTGGCCCTTCACCGATGTTGCT 1538
Db 243 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 259
```

RESULT 5

US-10-995-561-631
Sequence 631, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 631
LENGTH: 642
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-631

Alignment Scores:
Pred. No.: 3,11e-40 Length: 642
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

US-09-940-235-9 (1-1541) x US-10-995-561-631 (1-642)

```
QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGTGGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyrMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGGCGCAA---CAGATTGTACCATAGCTGAGAGTGTGTTGATCATCTGCTGGGACT 1247
Db 174 LysGlyGluThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
QY 1248 TCCTATGTGTGGGAGAAACGTGGGAGAACCCCTACCAAGCTGGAGTGGTAGATTGT 1307
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 213
QY 1308 ACTTGCTGGGAGAGCGGACGCATCATTGCATCTTAGAAATAGATGCAACGAT 1367
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGACACAAAGACATCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGTCCAGTGCATCTGCACAGCAACCGCCGAGGAGTGGAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGCAGACACATCGAGCGGATCTGGCCCTTCACCGATGTTGCT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290
```

RESULT 6

US-10-995-561-622
Sequence 622, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 622
LENGTH: 657
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-622

Alignment Scores:
Pred. No.: 3,12e-40 Length: 657
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

US-09-940-235-9 (1-1541) x US-10-995-561-622 (1-657)

Qy	1143	AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC	1190
Db	155	ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly	173
Qy	1191	CAGCGCAA---CAGATTGTACCATAGCTCAGAAAGTGTTTTGATCATGCTGCTGGGACT	1247
Db	174	LysGlyGluThrTrpCysLysProIleAlaGluLysCysPheAspHisAlaIaGlyThr	193
Qy	1248	TCCTATGTGTCGGAGAAACGTGGAGAACCCCTACCAAGCTCGATGATGGTAGATTGT	1307
Db	194	SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys	213
Qy	1308	ACTTGCCTGGAGAGCGGACGGACGCATCACTTGCATCTCTAGAAATAGATCAACGAT	1367
Db	214	ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp	233
Qy	1368	CAGGACACAAAGGACATCCTATAGAAATGGAGACACCTGGAGCAAGAAGGATAATCCAGGA	1427
Db	234	GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly	253
Qy	1428	AACCTGCTCCAGTCATCTGCACAGGCAACGCCGAGGAGAGTGGAAAGTGTGAGAGGCAC	1487
Db	254	AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluThrLysCysGluArgHis	273
Qy	1488	ACCTCTGTGCAGACCAATCGAGCGGATCTGCCCTTCCCGATGTTGCT	1538
Db	274	ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg	290
RESULT 7			
US-10-995-561-629			
; Sequence 629, Application US/10995561			
; Publication No. US20050272054A1			
; GENERAL INFORMATION:			
; APPLICANT: CARGILL, Michele et al.			
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH			
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF			
; TITLE OF INVENTION: DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001559			
; CURRENT APPLICATION NUMBER: US/10/995,561			
; CURRENT FILING DATE: 2004-11-24			
; NUMBER OF SEQ ID NOS: 85702			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 629			
; LENGTH: 984			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-995-561-629			
Alignment Scores:			
Pred. No.: 3,29e-40 Length: 984			
Score: 632.00 Matches: 117			
Percent Similarity: 87.59% Conservative: 3			
Best Local Similarity: 85.40% Mismatches: 11			
Query Match: 23.27% Indels: 6			
DB: 6 Gaps: 3			
US-09-940-235-9 (1-1541) x US-10-995-561-629 (1-984)			
Qy	1143	AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC	1190
Db	155	ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly	173
Qy	1191	CAGCGCAA---CAGATTGTACCATAGCTCAGAAAGTGTTTTGATCATGCTGCTGGGACT	1247
Db	174	LysGlyGluThrTrpCysLysProIleAlaGluLysCysPheAspHisAlaIaGlyThr	193
Qy	1248	TCCTATGTGTCGGAGAAACGTGGAGAACCCCTACCAAGCTCGATGATGGTAGATTGT	1307
Db	194	SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys	213
Qy	1308	ACTTGCCTGGAGAGCGGACGGACGCATCACTTGCATCTCTAGAAATAGATCAACGAT	1367

```

Db      214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
Qy      1368 CAGGACACACAGGACATCTATAGAAATGGAGACACCTGGAGCAAGAAGGATAATCGAGGA 1427
Db      234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerIlyAspAsnArgGly 253
Qy      1428 AACCTGCTCCAGTCATCTGCACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGCAC 1487
Db      254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpIlyCysGluArgHis 273
Qy      1488 ACCTCTGTGCAGACCAACATCGAGCGGATCTGGCCCTTACCCGATGTTGTT 1538
Db      274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 8
US-10-995-561-633
; Sequence 633, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 633
; LENGTH: 2296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-633

Alignment Scores:
Pred. No.: 3,68e-40 Length: 2296
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: 6 Gaps: 3

US-09-940-235-9 (1-1541) x US-10-995-561-633 (1-2296)
Qy      1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGTTGTTGGC 1190
Db      155 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 173
Qy      1191 CAGCGCGCA---CAGATTGTACCCATGCTGAGAAGTGTTTGATCATGCTGCTGGGACT 1247
Db      174 LysGlyGluTrpThrCysIysProIleAlaGluIlyCysPheAspHisAlaAlaGlyThr 193
Qy      1248 TCCTATGTTGTCGGAGAAAGCTGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGT 1307
Db      194 SerTyrValValGlyGluThrTrpGluIysProTyrGlnGlyTrpMetValAspCys 213
Qy      1308 ACTTGTCTGGAGAGCGAGCGGACGCATCCTGCATCTCTAGAATAGATCAACGAT 1367
Db      214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
Qy      1368 CAGGACACACAGGACATCTATAGAAATGGAGACACCTGGAGCAAGAAGGATAATCGAGGA 1427
Db      234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerIlyAspAsnArgGly 253
Qy      1428 AACCTGCTCCAGTCATCTGCACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGCAC 1487
Db      254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpIlyCysGluArgHis 273
Qy      1488 ACCTCTGTGCAGACCAACATCGAGCGGATCTGGCCCTTACCCGATGTTGTT 1538
Db      274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 9

```

RESULT 9

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-627

Alignment Scores:
Pred. No.:      3,7e-40          Length:      2355
Score:           632.00         Matches:     117
Percent Similarity: 87.59%       Conservative:   3
Best Local Similarity: 85.40%     Mismatches:    11
Query Match:      23.27%        Indels:        6
DB:               6              Gaps:          3

US-09-940-235-9 (1-1541) x US-10-995-561-627 (1-2355)
Qy 1143 AAGCGACCCGAAGGAGAAATGCTAGCTACCATTTA-----GCTGGTGTGGC 1190
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 155 ArgArgProHis---GluthrGlyGlyTy-MetLeuGluCysValCysLeuGlyAenGly 173
Qy 1191 CAGCGCAA---CAGATTGTACCATAGCTGAGAAGTTTTCATCATGCTGCTGGACT 1247
   :::::|||||::|||::|||::|||::|||::|||::|||::|||
Db 174 LyeGlyGluThrPrThrCysLysProIleAlaGluLysCysPheAspHisAlaLaGlyThr 193
Qy 1248 TCCTATCTGTGCGAGAAAAGTGGGAGAACCCCTTACCAAGCGCTGGATGATGTTAGATTGT 1307
   ::|||::|||::|||::|||::|||::|||::|||::|||
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 213
Qy 1308 ACTTGCTGGGAGAGGCAGCGACGCATCACCTTGCACTTCTAGAAATAGATCAACGAT 1367
   |||||::|||::|||::|||::|||::|||::|||::|||
Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThr-CysThrSerArgAsnA-gCysAsnAsp 233
Qy 1368 CAGGACACAGGACATCCTATAGAAATGGAGACACCTGGAGCAAGAGGATATCGAGGA 1427
   |||||::|||::|||::|||::|||::|||::|||::|||
Db 234 GlNAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnA-rgGly 253
Qy 1428 AACCTGCTCCAGTGCATCTCACAGGCCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCAC 1487
   |||||::|||::|||::|||::|||::|||::|||::|||
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyYargGlyGluTrpLysCysGluA-rgHis 273
Qy 1488 ACCTCTGTGAGACCACATCGACCGGATCTGGCCCCCTTCCACCGATGTTGCT 1538
   |||||::|||::|||::|||::|||::|||::|||::|||
Db 274 Thr-SerValGlnThrThrSerSerGlySerGlyProPheThrAspValA-rg 290

RESULT 11
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1545

Alignment Scores:
Pred. No.:      3,7e-40          Length:      2384
Score:           632.00         Matches:     117
Percent Similarity: 87.59%       Conservative:   3
Best Local Similarity: 85.40%     Mismatches:    11
Query Match:      23.27%        Indels:        6
DB:               6              Gaps:          3

```


US-09-940-235-9 (1-1541) x US-10-821-234-1545 (1-2384)

```

QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGC 1190
Db 184 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 202
QY 1191 CAGGCGCAA---CAGATTGTACCATAGCTGAGAGAGTGTGTGATCATGCTGGGACT 1247
Db 203 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 222
QY 1248 TCCTATGTGCTCGGAGAAAGCTGGAGAGCCCTACCAAGCTCGATGATGATGATGT 1307
Db 223 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 242
QY 1308 ACTTGCTGGGAGAGGACGAGCAGCATCATTGTCATCTTCAAAATAGATGCAACGAT 1367
Db 243 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 262
QY 1368 CAGGACACAAAGCATCTCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 263 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 282
QY 1428 AACCTGCTCAGTCATCTCCACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGCAC 1487
Db 283 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 302
QY 1488 ACTCTGTGTCAGACCATCGAGCGGATCTGGCCCTTCACCGATGTTTCTGT 1538
Db 303 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 319

```

RESULT 12

```

US-10-995-561-626
; Sequence 626, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 626
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-626

```

```

Alignment Scores:
Pred. No.: 3,7e-40 Length: 2386
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservativeness: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 23.27% Indels: 6
DB: Gaps: 3

```

US-09-940-235-9 (1-1541) x US-10-995-561-626 (1-2386)

```

QY 1143 AAGCGACCCGAGGAGAGATGCTAGCTACCATTTA-----GCTGGTGGTGC 1190
Db 155 ArgArgProHis---GluThrGlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGly 173
QY 1191 CAGGCGCAA---CAGATTGTACCATAGCTGAGAGAGTGTGTGATCATGCTGGGACT 1247
Db 174 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
QY 1248 TCCTATGTGCTCGGAGAAAGCTGGAGAGCCCTACCAAGCTCGATGATGATGATGT 1307
Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 213
QY 1308 ACTTGCTGGGAGAGGACGAGCAGCATCATTGTCATCTTCAAAATAGATGCAACGAT 1367

```

```

Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233
QY 1368 CAGGACACAAAGCATCTCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427
Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253
QY 1428 AACCTGCTCAGTCATCTGCACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGCAC 1487
Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
QY 1488 ACCTCTGTGTCAGACCATCGAGCGGATCTGGCCCTTCACCGATGTTTCTGT 1538
Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

```

RESULT 13

```

US-10-995-561-621
; Sequence 621, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-621

```

```

Alignment Scores:
Pred. No.: 8e-13 Length: 1341
Score: 265.50 Matches: 118
Percent Similarity: 33.90% Conservativeness: 62
Best Local Similarity: 22.22% Mismatches: 160
Query Match: 9.78% Indels: 191
DB: Gaps: 23

```

US-09-940-235-9 (1-1541) x US-10-995-561-621 (1-1341)

```

QY 72 AATAACAGCCAAATGGTGTGTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATT 131
Db 848 AspSerSerValValValSer-----GlyLeuMetValAlaThrLysTyrGluVal 865
QY 132 AGTCTTAAATTTTTGAAATCGATCTAAACATCAGACCTGCTCATGCGGAGGAAACAGAG 191
Db 866 SerValTyrAlaLeuLysAspThrLeuThrSerArgProAlaGlnGlyValValThr 885
QY 192 ---CAAGGCTTAAGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCACAT 248
Db 886 LeuGluAsnValSerProArgAlaArgValThrAsp----- 899
QY 249 AAAGCTTGAGAAAGCTGACTTAAAGGCTATTCAAGAACAAATTCATGCTGCTAACGCTCCAC 308
Db 899 ----- 899
QY 309 AGTAACGACGACTACTTTGAGGTCAATGATTTTCAAGCGGATGCAACATTAATCTGATCGA 368
Db 900 -----AlaThrGluThrThrIleThrIle--- 907
QY 369 AACGGCAAGGTCTACTTTGCTGCAAAAGATGGTTGGTAACTTGCCTGCGGACCAACCTGTC 428
Db 908 -----SerTrpArgThrLysThrGluThrIle 916
QY 429 CAAGAAATTTTGTAAAGCGGACATGTGCGCGTTAGACCATATAAGAAACCAATACAA 488
Db 917 ThrGlyPheGlnValAspAla-----ValProAlaAsnGlyGlnThrProIleGln 933

```

```
QY 489 AACCAAGCGAATCTGTTGATGTGGAATATACGTACAGTTTACTCCCTTAAACCCGTGAT 548
Db 934 ArgThrIleLysProAspValArgSerTyrThrIle-----ThrGlyLeuGlnProGly 951
QY 549 GACGATTTTCAGACCGAGTCTCAAGATATACGTATTTGAAACACTAGCTATCGGTGAC 608
Db 952 ThrAspTyrLys-----IleTyrLeuTyrThrLeuAsnAsp 963
QY 609 ACCATCACATCTCAAGAATTACTAGTCAAGCA----- 641
Db 964 AsnAlaArgSerSerProValValIleAspAlaSerThrAlaIleAspAlaProSerAsn 983
QY 642 -----CAAAGCATTTTAAACAAAAACCAACCA----- 668
Db 984 LeuArgPheLeuAlaThrThrProAsnSerLeuLeuValSerTrpGlnProProArgAla 1003
QY 669 -----GGCTATACGATT-----TATGAACGTGACTCTCAATCGTCACTCATGACAAT 716
Db 1004 ArgIleThrGlyTyrIleIleLysTyrGluLysProGlySerProPro----- 1019
QY 717 GACATTTTCCTGACGATTTTACCAATGGATCAA----- 749
Db 1020 -----ArgGluValValProArgProArgProGlyValThrGluAlaThrIleThr 1036
QY 750 -----GAGTTTACTTACCGTGT-----AAAAATCGGGAACAA 782
Db 1037 GlyLeuGluProGlyThrGluTyrThrIleTyrValIleAlaLeuLysAsnGlnLys 1056
QY 783 GCTTATAGATC-----AATAAAAATCTGGTCTGAATGAAGAATAAACAACACTGAC 836
Db 1057 SerGluProLeuIleGlyArgLysLysThrGly---GlnGluAlaLeuSerGlnThrThr 1075
QY 837 CTGATCTCTGAGAAATATTACGTCCTTAAAGGGGAAAAAGCGGTATGATCCCTTTGAT 896
Db 1076 Ile-----SerTrpAlaProPheGln 1082
QY 897 CGCAGTCACCTGAAACTGTTCCACCAATAACGTGATGTCGTATACCAACGAATTGCTA 956
Db 1083 -----AspThrSerGluTyrIle 1088
QY 957 AAAAGTGACGACCTTTAAACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGAT 1016
Db 1089 IleSerCysHisProValGlyThrAspGluGluProLeuGlnPheArg-----Val 1105
QY 1017 CCTCGTATAGGCTAAACTACTCTACAACTCTCGATCGTCTTTGGTATATGAGCATAT 1076
Db 1106 ProGlyThrSerThrSerAla----- 1112
QY 1077 ACCTTAACCTGGA-----AAAGTA 1094
Db 1113 ThrLeuThrGlyLeuThrArgGlyAlaThrTyrAsnIleValGluAlaLeuLysAsp 1132
QY 1095 GAGGATAATCAGATGACACCAACCGTATCATACACCGTTTATATGGCGAAGCAGCCGAA 1154
Db 1133 GlnGlnArgHisLysValArgGluGluValValThrValGlyAsnSerValAsn---Glu 1151
QY 1155 GGAGAGAATCTAGTACCAATTTAGTGTGTGGCGCCAGCGCAACAGATTGTACCCATA 1214
Db 1152 GlyLeuAsnGln-----ProThr 1157
QY 1215 GCTGAGAAGTGTTTTGATCATGCTGCGGACTTCCTATCTGTCGCGAAGAACGTGGAG 1274
Db 1158 AspAspSerCysPheAspProTyrThrValSerHisTyrAlaValGlyAspGluTrpGlu 1177
QY 1275 AAGCCCTACCAAGGTGGATGATGTAGATTGTACTTGTCTGGAGAAGCGCAGCGACGC 1334
Db 1178 ArgMetSerGluSerGlyPheLysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHis 1197
QY 1335 ATCATTCTGACTTCTAGAAATATAGATGCAACGATCAGGACACAGAGGACATCTATAGAAAT 1394
Db 1198 PheArgCysAspSerSerArgTrpCysHisAspAsnGly-----ValAsnTyrLysIle 1215
QY 1395 GGAGACACTGTGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1454
```

```
Db 1216 GlyLeuLysTrpAspArgGlnGlyGluAsnGlyGlnMetMetSerCysThrCysLeuGly 1235
QY 1455 AACGGCCGAGAGAGTGAAGTCTGAGAGGCAC 1487
Db 1236 AsnGlyLysGlyGluPheLysCysAspProHis 1246
RESULT 14
US-10-995-561-632
; Sequence 632, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 632
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-632
Alignment Scores:
Pred. No.: 1,23e-12 Length: 693
Score: 262.50 Matches: 123
Percent Similarity: 32.40% Conservative: 62
Best Local Similarity: 21.54% Mismatches: 179
Query Match: 9.66% Indels: 207
DB: 6 Gaps: 23
US-09-940-235-9 (1-1541) x US-10-995-561-632 (1-693)
QY 72 AATAACAGCAATTTGTTGTAGCGTGTCTGTTGAGGGGAGCAATCAAGACATT 131
Db 136 AspSerSerValValValSer-----GlyLeuMetValAlaThrLysTyrGluVal 153
QY 132 AGTCTTAAATTTTGAATCGATCAATCAATCAGACCTGCTCATGAGGAGAAACAGAG 191
Db 154 SerValTyrAlaLeuLysAspThrLeuThrSerArgProAlaGlnGlyValValThrThr 173
QY 192 ---CAGGCTTAAGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAAT 248
Db 174 LeuGluAsnValSerProArgArgAlaArgValThrAsp----- 187
QY 249 AAATTGAGAAAGCTGACTTACTTAAGGCTATTCAAGAACAAATTGATCGCTAACGTCAC 308
Db 187 ----- 187
QY 309 AGTAACAGCAGCTACTTTTGAGGTTCATTGATTTTGAAGCGATGCAACCACTTACTGATCGA 368
Db 188 -----AlaThrGluThrThrIleThrIle--- 195
QY 369 AACGGCAAGGTCTACTTTGCTGCAAAAGATGGTTCGGTAAACCTTGGCCGACCAACCTGTC 428
Db 196 -----SerTrpArgThrLysThrGluThrIle 204
QY 429 CAAGAAATTTTGTAAAGCGACATGTGCGGTGTAGACCATATAAGAAACCAATACAA 488
Db 205 ThrGlyPheGlnValAspAla-----ValProAlaAsnGlyGlnThrProIleGln 221
QY 489 AACCAAGCAAAATCTGTTGATGTGGAATATACGTATGTCAGTTTACTCCTTAAACCCGTGAT 548
Db 222 ArgThrIleLysProAspValArgSerTyrThrIle-----ThrGlyLeuGlnProGly 239
QY 549 GACGATTTTCAGACCGAGTCTCAAGATATACGTATTTGAAACACTAGCTATCGGTGAC 608
Db 240 ThrAspTyrLys-----IleTyrLeuTyrThrLeuAsnAsp 251
```

```
QY 609 ACCATCACATCTCAAGAAATTACTAGCTCAAGCA----- 641
Db 252 AsnAlaArgSerProValValIleAspAlaSerThrAlaIleAspAlaProSerAsn 271
QY 642 -----CAAGCATTTTAAACAAAACCACCCA----- 668
Db 272 LeuArgPheLeuAlaThrProAsnSerLeuLeuValSerTrpGlnProProArgAla 291
QY 669 -----GGCTATACGATT-----TATGACGCTGACTCCTCAATCGTCACTCATGACAAAT 716
Db 292 ArgIleThrGlyTyrllelleleTyrlleGluLysProGlySerProPro----- 307
QY 717 GACATTTCCGTAGGATTTACCAATGATCAAA----- 749
Db 308 -----ArgGluValValProArgProArgProGlyValThrGluAlaThrIleThr 324
QY 750 -----GAGTTTACTTACCGTGTAAATAATCGGAAACGACTTATAGG 791
Db 325 GlyLeuGluProGlyThrGluTyrlleThrIleTyrlleVal-----IleAlaLeuLys 340
QY 792 ATCAATAAAAAATCT-----GGTCTGAATGAAGAAATAAAACAACACTGACCTG 839
Db 341 AsnAsnGlnLysSerGluProLeuIlelleGlyArgLysThrValGlnLysThrProPhe 360
QY 840 ATCTCTGAGAAATATTACGTCCTTAAANAAGG----- 872
Db 361 ValThrHisProGlyTyrlleThrAspThrGlyAsnGlylleGlnLeuProGlyThrSerGlyGln 380
QY 872 ----- 872
Db 381 GlnProSerValGlyGlnMetIlePheGluGluHisGlyPheArgThrThrPro 400
QY 873 -----GAAAGCGGTATGATCCC-----TTTGATCGC 899
Db 401 ProThrThrAlaThrProIleArgHisArgProArgProTyrlleProAsnValGlyGln 420
QY 900 AGTCACTGAAACTGTTCCACCATCAATACGTTGATGTC---GATACCAACGAATTCGTA 956
Db 421 GluAlaLeuSerGlnThrThrIleSerTrpAlaProPheGlnAspThrSerGluTyrlle 440
QY 957 AAAGTGGAGCAGCTTTAACAGCTAGCGAAGCTTACACTTACAGTCTCAGAGATTTATACGAT 1016
Db 441 IleSerCysHisProValGlyThrAspGluGluProLeuGlnPheArg-----Val 457
QY 1017 CCTCGTGATAAGCTAAACTACTCTACAACTCTGATCGATCGTCTTGTATATGAGCATAT 1076
Db 458 ProGlyThrSerThrSerAla----- 464
QY 1077 ACCTTAACCTGGA-----AAAGTA 1094
Db 465 ThrLeuThrGlyLeuThrArgGlyAlaThrTyrlleValGluAlaLeuLysAsp 484
QY 1095 GAGGATAAATCAGCATGACACCAACCGTATCATCAACCGTTTATATGGCAAGCGACCGGAA 1154
Db 485 GlnGlnArgHisLysValArgGluGluValValThrValGlyAsnSerValAsn---Glu 503
QY 1155 GGAGAGAATGCTAGTACCACTTATAGTGTGGTGGCCAGCGCAACAGATTGACCCATA 1214
Db 504 GlyLeuAsnGln-----ProThr 509
QY 1215 GCTGAGAAGTGTTCATCATGCTGGGACTTCCTATGCTGGGAGAAACGTCGGAG 1274
Db 510 AspAspSerCysPheAspProTyrlleThrValSerHisTyrlleValGlyAspGluTrpGlu 529
QY 1275 AAGCCCTACCAAGCTGGATGATGTGTACTTGGCTGGGAGAAGCGACGGACGC 1334
Db 530 ArgMetSerGluSerGlyPheLysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHis 549
QY 1335 ATCACTTGCACTTCTAGAATAATAGATCAACAGATCAGGACACAAGGACATCTATAGAAAT 1394
Db 550 PheArgCysAspSerSerArgTTrpCysHisAspAsnGly-----ValAsnTyrlle 567
QY 1395 GGAGACACCTGGAGCAAGAGGATATATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1454
```

```
Db 568 GlyGluLysTrpAspArgGlnGlyGluAsnGlyGlnMetSerCysThrCysLeuGly 587
QY 1455 AACGCCGAGGAGAGTGGAGTGTGAGAGGCAC 1487
Db 588 AsnGlyLysGlyGluPheLysCysAspProHis 598

RESULT 15
US-10-995-561-630
; Sequence 630, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 630
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-630

Alignment Scores:
Pred. No.: 1,34e-12 Length: 1315
Score: 262.50 Matches: 123
Percent Similarity: 32.40% Conservative: 62
Best Local Similarity: 21.54% Mismatches: 179
Query Match: 9.66% Indels: 207
DB: 6 Gaps: 23

US-09-940-235-9 (1-1541) x US-10-995-561-630 (1-1315)
QY 72 AATPAACGCCAATTTGGTTGTTAGCGTGTCTGTTAGGAGGAGCAATCAAGACATT 131
Db 758 AspSerSerValValValSer-----GlyLeuMetValAlaThrLysTyrlleVal 775
QY 132 AGTCTTAAATTTTGAATTCGATCTAACATCAGACCTGCTATGAGGAGGAGCAAGACAG 191
Db 776 SerValTyrlleLysAspThrLeuThrSerArgProAlaGlnGlyValValThrThr 795
QY 192 ---CAAGCTTAAGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCATGTGCAT 248
Db 796 LeuGluAsnValSerProArgArgAlaArgValThrAsp----- 809
QY 249 AAACCTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGACAAATTGATCGTAAACGTCAC 308
Db 809 ----- 809
QY 309 AGTAACGACGACTACTTTGAGGTCAATTGTTTTCGACCGCATGCAACCAATTACTGATCGA 368
Db 810 -----AlaThrGluThrThrIleThrIle--- 817
QY 369 AACGGCAAGGTCTACTTTGCTGCAAAAGATGTTTCGGTAACTTTCGCGAGCCCAACTGTC 428
Db 818 -----SerTrpArgThrLysThrGluThrIle 826
QY 429 CAAGAAATTTTGTGAACGGACATGTGCGGTATAGACCATATAAAGAAACCAATACAA 488
Db 827 ThrGlyPheGlnValAspAla-----ValProAlaAsnGlyGlnThrProIleGln 843
QY 489 AACCAACGCAAACTGTTGATGTGGAATATATCTGATACAGTTTACTCCCTTAACCCCTGAT 548
Db 844 ArgThrIleLysProAspValArgSerTyrlle-----ThrGlyLeuGlnProGly 861
QY 549 GACGATTTCAAGCCAGGTCTCAAAAGATCTAAAGTATTGAAACACACTAGCTATCGGTGAC 608
Db 862 ThrAspTyrlleLys-----IleTyrlleTyrlleThrLeuAsnAsp 873
```

Qy 609 ACCATCATCTCAAGAAATTACTAGCTCAAGCA----- 641
 Db 874 AsnAlaArgSerProValValIleAspAlaSerThrAlaIleAspAlaProSerAsn 893
 Qy 642 -----CAAAGCATTTAAACAAACCAACCCCA----- 668
 Db 894 LeuArgPheLeuAlaThrThrProAsnSerLeuLeuValSerTrpGlnProProArgAla 913
 Qy 669 -----GGCTATACGATT---TATGAACGTGACTCTCAATCCTCATCTCATGACAAAT 716
 Db 914 ArgIleThrGlyTyrIleIleLysTyrGluLysProGlySerProPro----- 929
 Qy 717 GACATTTTCCGTACGATTTTACCAATGGATCAA----- 749
 Db 930 -----ArgGluValValProArgProArgProGlyValThrGluAlaThrIleThr 946
 Qy 750 -----GAGTTTACTACCGTGTAAAAATCGGGAACAAGCTTATAGG 791
 Db 947 GlyLeuGluProGlyThrGluTyrThrIleTyrVal-----IleAlaLeuLys 962
 Qy 792 ATCAATAAAATCT-----GGTCTGAATGAAGAAATAAACAACACTGACCTG 839
 Db 963 AsnAsnGlnLysSerGluProLeuIleGlyArgLysThrValGlnLysThrProPhe 982
 Qy 840 ATCTCTGAGAAATATTACGTCCTTAAAAAAGG----- 872
 Db 983 ValThrHisProGlyTyrAspThrGlyAsnGlyIleGlnLeuProGlyThrSerGlyGln 1002
 Qy 872 ----- 872
 Db 1003 GlnProSerValGlyGlnGlnMetIlePheGluGluHisGlyPheArgArgThrThrPro 1022
 Qy 873 -----GAAAAGCCGTATGATCCC---TTTGATCCG 899
 Db 1023 ProThrThrAlaThrProIleArgHisArgProArgProTyrProProAsnValGlyGln 1042
 Qy 900 AGTCACCTTGAAACTGTTCAACCATCAAAATAGCTTGATGTC---GATACCAACGAATTGCTA 956
 Db 1043 GluAlaLeuSerGlnThrThrIleSerTrpAlaProPheGlnAspThrSerGluTyrIle 1062
 Qy 957 AAAAGTGACAGCTCTTAACAGCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGAT 1016
 Db 1063 IleSerCysHisProValGlyThrAspGluGluProLeuGlnPheArg-----Val 1079
 Qy 1017 CCTCGTGATAGGCTAAACTCTCTACAAACAATCTCGATGCTTTGGTATTATGGAATAT 1076
 Db 1080 ProGlyThrSerThrSerAla----- 1086
 Qy 1077 ACCTTAACCTGGA-----AAAGTA 1094
 Db 1087 ThrLeuThrGlyLeuThrArgGlyAlaThrTyrAsnIleIleValGluAlaLeuLysAsp 1106
 Qy 1095 GAGGATAATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAA 1154
 Db 1107 GlnGlnArgHisLysValArgGluGluValValThrValGlyAsnSerValAsn---Glu 1125
 Qy 1155 GGAGAGATGCTAGCTACCATTTAGCTGTGTGTGGCCGCGCAACAGATTGTACCCATA 1214
 Db 1126 GlyLeuAsnGln-----ProThr 1131
 Qy 1215 GCTGAGAAGTGTGTTTGATCATCTGCTGGGACTTCTTATGTGTGGAGAAACGTGGGAG 1274
 Db 1132 AspAspSerCysPheAspProTyrThrValSerHisTyrAlaValGlyAspGluTrpGlu 1151
 Qy 1275 AAGCCCTACCAAGGCTGGATGATGATTGCTTGTGGGAGAGCGACGCGACGC 1334
 Db 1152 ArgMetSerGluSerGlyPheLysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHis 1171
 Qy 1335 ATCACTTCGACTCTAGAAAATAGATCAACAGATCAGGACACAAGGACATCTCTATAGATT 1394
 Db 1172 PheArgCysAspSerSerArgTrpCysHisAspAsnGly-----ValAsnTyrLysIle 1189
 Qy 1395 GGAGACACTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1454

Db 1190 GlyGluLysTrpAspArgGlnGlyGluAsnGlyGlnMetMetSerCysThrCysLeuGly 1209
 Qy 1455 AACGCCCGAGGAGAGTGGAAAGTGTGAGAGGCAC 1487
 Db 1210 AsnGlyLysGlyGluPheLysCysAspProHis 1220

Search completed: January 28, 2006, 02:57:30
 Job time : 25.4619 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: January 28, 2006, 01:49:35 ; Search time 23.6156 Seconds
(without alignments)
12556.953 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 2716
Sequence: 1 ttgtttaacttaagaagg.....ccttcaccgatgttcgtag 1541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xl
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_27012006_144218_27578/app_query.fasta_1.7708
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1_1_185 @runat_27012006_144218_27578 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	73.0	440	1	streptokinase prec
2	1951.5	71.9	415	1	streptokinase (EC
3	1942	71.5	440	2	streptokinase G pr
4	1779	65.5	440	2	streptokinase A pr
5	1673	61.6	414	2	streptokinase - St
6	1670	61.5	440	2	streptokinase A pr
7	667	24.6	197	2	streptokinase - St
8	632	23.3	2386	1	fibronectin precu
9	620	22.8	128	2	streptokinase A (E
10	610	22.5	2265	1	fibronectin - bovi
11	585.5	21.6	2477	2	fibronectin precu
12	569	20.9	128	2	streptokinase A (E
13	557	20.5	128	2	streptokinase A (E
14	552	20.3	128	2	streptokinase A (E

15	544	20.0	2481	2	A43908	fibronectin - Afri
16	423	15.6	128	2	S77676	streptokinase A (E
17	419	15.4	128	2	S77673	streptokinase A (E
18	419	15.4	128	2	S77687	streptokinase A (E
19	417	15.4	128	2	S77682	streptokinase A (E
20	416	15.3	128	2	S77678	streptokinase A (E
21	408	15.0	128	2	S77677	streptokinase A (E
22	407	15.0	128	2	S77685	streptokinase A (E
23	401	14.8	128	2	S77683	streptokinase A (E
24	398	14.7	128	2	S77686	streptokinase A (E
25	397	14.6	128	2	S77684	streptokinase A (E
26	396	14.6	128	2	S77681	streptokinase A (E
27	390	14.4	128	2	S77674	streptokinase A (E
28	386	14.2	128	2	S77675	streptokinase A (E
29	221	8.1	1020	2	A29355	fibronectin - chic
30	182	6.7	190	2	151279	fibronectin - east
31	146.5	5.4	2231	2	D71870	hypothetical prote
32	130	4.8	1745	2	A46431	tight junction-ass
33	130	4.8	2334	2	S32920	cell wall-associat
34	129.5	4.8	940	2	AD1374	internalin protein
35	126	4.6	4688	2	F82885	hypothetical prote
36	125.5	4.6	940	2	AB1744	internalin protein
37	125	4.6	1220	1	DU8EC3	DNA-directed DNA p
38	124.5	4.6	627	2	F84194	Htr14 transducer
39	124.5	4.6	1315	2	T28679	fibrinogen-binding
40	123	4.5	1019	2	T50251	hypothetical coile
41	122.5	4.5	1435	2	S54697	DNA polymerase III
42	122.5	4.5	1435	2	C90596	hypothetical prote
43	121	4.5	785	2	D71653	cell surface antig
44	121	4.5	1237	2	AC1583	internalin protein
45	121	4.5	1385	2	D89824	hypothetical prote

ALIGNMENTS

RESULT 1

A22801
streptokinase precursor - Streptococcus "equisimilis"
C;Species: Streptococcus "equisimilis"
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A22801
R;Malke, H.; Roe, B.; Ferretti, J.J.
Gene 34, 357-362, 1985
A;Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H
A;Reference number: A22801; MUID:85232082; PMID:2989113
A;Accession: A22801
A;Molecule type: DNA
A;Residues: 1-440 <MAL>
A;Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:g407876; PIDN
A;Experimental source: strain H46A
C;Genetics:
A;Gene: skc
C;Superfamily: streptokinase

Alignment Scores:
Pred. No.: 2.53e-127 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.05% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-9 (1-1541) x A22801 (1-440)

QY	33	ATAGCTGCTCCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCAATTCGTTGTT	92
DB	27	IIeAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	46
QY	93	AGCGTTGCTGGTACTGTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC	152
DB	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLeysPhePheGluIle	66
QY	153	GATCTAATCATCAGCACCTGCTCATGTGAGGAGAAAGACAGACAGAGCTTAAGTCCAAATCA	212

Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCATTTGCTACTGATAGTCGCGCATGTCCACATAAATCTGAGAAAGCTGCTACTTA 272
Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLysLeuGlnLysAlaAspLeuLeu 106
QY 273 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCCTCACAGTAACGACCACTACTTTGAGGTC 332
Db 107 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
QY 333 ATTGATTTTGCAGCGATGCAACCAATTAATGATCGAAACGCGCAAGTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAGATGCTTCGGTAACTTCGCGACCCCAACCTGTCCAAGATTTTGTGTAACGCGACAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 453 GTGCGCGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 167 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCTCAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACCATCACATCTCAAGATTACTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACATGACATGACATTTTCGTACGATTTTACCATGGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATAAAAAATCTGTCGTG 812
Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
QY 813 AATGAGAAATAAACAACACTGATCTCTGAGAAATATTACGTCTTAAAAAGGG 872
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
QY 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACATTGAAACTCTTCAACCATCAAAATACGTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGCGNATCAACGAATTCATAAAGTAGCAGCTCTTAACAGCTAGCGAACGTAAC 992
Db 327 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTCAGAGTTTATACGATCCTCGTGAAGGCTAAACTACTCTACACATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGTGATATTATGCACTATACCTTAATCGGAAAAGTAGAGGATATCAACGATGAC 1112
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1113 ACCAACCGTATCATAAACCGTTTATATGGCGAAGCGACCCGAGGAGAGATGCTAGCTAC 1172
Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGlyGluAsnAlaSerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409

RESULT 2
BZSO

streptokinase (EC 3.4.-.-) - Streptococcus sp.
C;Species: Streptococcus sp.
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C;Accession: A00967
R;Jackson, K.W.; Tang, J.
Biochemistry 21, 6620-6625, 1982
A;Title: Complete amino acid sequence of streptokinase and its homology with serine protease
A;Reference number: A00967; MUID:83127125; PMID:6760891
A;Accession: A00967
A;Molecule type: protein
A;Residues: 1-415 <JAC>
A;Cross-references: UNIPROT:P00779; UNIPARC:UPI0000172BF5
A;Note: 169-Asp and 181-Asp were also found
A;Note: this protein is not a protease, but it activates plasminogen by complexing with
C;Superfamily: streptokinase
C;Keywords: hydrolase

Alignment Scores: Pred. No.: 4,09e-125 Length: 415
Score: 1951.50 Matches: 380
Percent Similarity: 99.22% Conservative: 1
Best Local Similarity: 98.96% Mismatches: 2
Query Match: 71.85% Indels: 1
DB: Gaps: 1
US-09-940-235-9 (1-1541) x BZSO (1-415)

QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTTCTGTAATAATACAGCCAAATGGTTGTT 92
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 93 AGCGTGTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATATTTTGAATC 152
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 153 GATCTAACATCAAGACCTGCTCATGGAGGAGAAACAGACAGCAAGGCTTAAGTCCAAATCA 212
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACATTTGCTACTGATGATGCGCGATGTACATAAACTTGAGAAAGCTGCTTACTA 272
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLysLeuGluLysAlaAspLeuLeu 80
QY 273 AAGGCTATTCAAGAACAAATTGATCGTCAAGTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 81 LysAlaIleGlnGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 333 ATTGATTTTGCAGCGATGCAACCAATTAATGATCGAAGGCTTACTTTGCTGAC 392
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAAGATGCTTCGGTAACTTCGCGACCCCAACCTGTCCAAAGATTTTGTGTAACGCGACAT 452
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGCGCGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATACGTATCAGATTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 632
Db 181 LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATAGGATTTATGACGTCAC 692
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCATGGATCAAGAG 752

Db	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
Qy	753	TTTACTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	241	PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	260
Qy	813	AATGAAGAATAAACACACTGACCTGATCTCTGAGAAATAT---TACGTCCTTAAAAA	869
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluTyrLysTyrValLeuLysLys	280
Qy	870	GGGGAAGACCGTATCATCCCTTTGATCGGAGTCACCTTGAACCTGTTCCACCAATAC	929
Db	281	GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr	300
Qy	930	GTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGA	989
Db	301	ValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg	320
Qy	990	AACTTAGCTTCAGAGATTATACGATCTCTCGTGATAGGCTAAACTTACTCTACAACAT	1049
Db	321	AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn	340
Qy	1050	CTCGATGCTTTGGTATTATGGACTATACCTTAACCTGGAAGTACAGGATAATCAGAT	1109
Db	341	LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp	360
Qy	1110	GACACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAAGGAGAGATGCTAGC	1169
Db	361	AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer	380
Qy	1170	TACCAATTAGCT 1181	
Db	381	TyrHisLeuAla 384	
RESULT 3			
S02723			
streptokinase G precursor - Streptococcus sp. (strain 19908)			
C:Species: Streptococcus sp.			
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004			
C:Accession: S02723			
R:Walter, F.; Siegel, M.; Malke, H.			
Nucleic Acids Res. 17, 1262, 1989			
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.			
A:Reference number: S02723; MUID:89160265; PMID:2922269			
A:Accession: S02723			
A:Molecule type: DNA			
A:Residues: 1-440 <WAL>			
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:g47095; PIDN			
C:Genetics:			
A:Gene: skg			
C:Superfamily: streptokinase			
F:1-26/Domain: signal sequence #status predicted <SIG>			
F:27-440/Product: streptokinase #status predicted <MAT>			
Alignment Scores:			
Pred. No.:	1-82e-124	Length:	440
Score:	1942.00	Matches:	375
Percent Similarity:	98.43%	Conservative:	2
Best Local Similarity:	97.91%	Mismatches:	6
Query Match:	71.50%	Indels:	0
DB:	2	Gaps:	0
US-09-940-235-9 (1-1541) x S02723 (1-440)			
Qy	33	ATAGCTGGTCTCTGAATGGCTACTAGATCGTCTCTCTGTAAATAACCAACCAATTTGGTTGT	92
Db	27	IleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	46
Qy	93	ACGGTGTCTGGTACTGTGGGGGACGAATCAAGACATTAAGATCTTAAATTTTGTAAATC	152
Db	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	66
Qy	153	GATCTAACATCAGCACCTGCTCATCGGAGGAAAGACAGACGAGGCTTAAAGTCCAAATCA	212

Db	67	AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer	86
Qy	213	AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA	272
Db	87	LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106
Qy	273	AAGCTATTCAAGAACAAATTCATCGCTAAACCTCCACAGTACAGCAGCTACTTTGAGTGC	332
Db	107	LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	126
Qy	333	ATTGATTTTGCACGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTCTGAC	392
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	146
Qy	393	AAAGATGGTTCGGTAAACCTTGCAGCCCAACCTGTCGAAGAATTTTTCGTAAGCGGACAT	452
Db	147	LysAspGlySerValThrLeuProIleGlnProValGlnGluPheLeuLeuLysGlyHis	166
Qy	453	GTGCGCTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG	512
Db	167	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186
Qy	513	GAATPATCTGTACAGTTTACTCCCTTAAACCTCGATGACGATTTTCAGACCAAGTCTCAAA	572
Db	187	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys	206
Qy	573	GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAATTACTA	632
Db	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu	226
Qy	633	GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCAGCTATACGATTTATGAACGTGAC	692
Db	227	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	246
Qy	693	TCCTCAATCGCTACTCATGACAAATGACATTTTCCGTACGATTTTACCATAAGTCAAGAG	752
Db	247	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	266
Qy	753	TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	812
Db	267	PheThrTyrHisValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	286
Qy	813	AATGAGAAATAACACACTGACCTGATCTCTGAGAATATTTACGTCCTTAAAAAGGG	872
Db	287	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	306
Qy	873	GAAGAAGCGGTATCATCCCTTTGATCGCAGTCACCTTGAAACTGTTCCACCATCAAAATACGTT	932
Db	307	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326
Qy	933	GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCAACGTAAC	992
Db	327	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346
Qy	993	TTAGACTTCAGACATTTATACGATCTCGGTGATAAGCTAAACTACTCTTACAACATCTC	1052
Db	347	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	366
Qy	1053	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATAATCAGATGAC	1112
Db	367	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	386
Qy	1113	ACCAACCGTATACAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGATAATGCTAGCTAC	1172
Db	387	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	406
Qy	1173	CATTAGCT 1181	
Db	407	HisLeuAla 409	
RESULT 4			
S02724			

streptokinase A precursor - Streptococcus pyogenes (strain SF130/13)
C;Species: Streptococcus pyogenes
C;Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 09-Jul-2004
C;Accession: S02724
R;Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1261, 1989
A;Title: Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes typh
A;Reference number: S02724; MUID:89160264; PMID:2646590
A;Accession: S02724
A;Molecule type: DNA
A;Residues: 1-440 <WAL>
A;Cross-references: UNIPROT:P10520; UNIPARC:UPI000017021D; EMBL:X13399; NID:947435; PIDN
C;Genetics:
A;Gene: ska
C;Superfamily: streptokinase
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-440/Product: streptokinase #status predicted <MAT>

Alignment Scores:
Pred. No.: 2,21e-113 Length: 440
Score: 1779.00 Matches: 338
Percent Similarity: 93.73% Conservative: 21
Best Local Similarity: 88.25% Mismatches: 24
Query Match: 65.50% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S02724 (1-440)

QY 33 ATAGCTGTCCTGAATGGCTACTAGATCGTCTCTTCTGTAATAACAGCCAAATGGTGT 92
DB 27 IleAlaGlyTyrGlyTrpLeuProAspArgProIleAsnSerGlnLeuValVal 46
QY 93 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACATAGTCTTAAATTTTGAATC 152
DB 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
QY 153 GATTAACATCACGACCTCTCATGGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 212
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACATTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 272
DB 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AAGCGTATTCAGAACATTTGATCGCTAAGCTCCACAGTAAACGACACTACTTTGAGTTC 332
DB 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTCGAGCGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 392
DB 127 IleaspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAGATGTTTCGGTAACCTTGCCGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 452
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuValLysGlyHis 166
QY 453 GTGGCGCTAGACCATATAAGAAACCAACATACAAACCAACGAGCAAACTGTGTATGTCG 512
DB 167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACGAGTCTCAAA 572
DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 632
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAGACAAAGCATTTTAAACAAACACACCCAGGCTATACGATTTATGAACTGAC 692
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAATAATCGGACACAGCTTATAGGATCAATAAATAATCTGCTGTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY 813 AATGAAGAAATAAACACACACTGACTCTCTGAGAAATATTAGCTCTTAAAAAAGGG 872
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
QY 873 GAAAGCCGATGATCCCTTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAAAATACGTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCGATACCAACGAATTGCTTAAAGTCAAGAGCTCTTAAACAGTACGAGCAAGCTTAAC 992
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluGlyAsn 346
QY 993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAACTACTCTACAAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTCGAAAAAGTACAGGATAATCAGCATGAC 1112
Db 367 AspAlaPheAspIleMetAsnTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACGCTATCATACCGTTTATATGGGCAAGGACCCGAGGAGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409

RESULT 5
JU0292
streptokinase - Streptococcus pyogenes
C;Species: Streptococcus pyogenes
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: A43867; JU0292
R;Ohkuni, H.; Todome, Y.; Suzuki, H.; Mizuse, M.; Kotani, N.; Horiuchi, K.; Shikama, N.
Infect. Immun. 60, 278-283, 1992
A;Title: Immunochemical studies and complete amino acid sequence of the streptokinase
A;Reference number: A43867; MUID:92104686; PMID:1370275
A;Molecule type: protein
A;Accession: A43867
A;Residues: 1-414 <OHK>
A;Cross-references: UNIPROT:Q57391; UNIPARC:UPI0000175C82
A;Experimental source: M type 12 strain A374
A;Note: sequence extracted from NCBI backbone (NCBIP:74592)
C;Superfamily: streptokinase

Alignment Scores:
Pred. No.: 3,56e-106 Length: 414
Score: 1673.00 Matches: 324
Percent Similarity: 90.60% Conservative: 23
Best Local Similarity: 84.60% Mismatches: 36
Query Match: 61.60% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x JU0292 (1-414)

QY 33 ATAGCTGGTCTGAATGGCTACTAGATCGTCTCTTCTGTAATAACAGCCAAATGGTGT 92
Db 1 IleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 93 AGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAATC 152
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnGluIleSerLeuLysPhePheGluIle 40
QY 153 GATCTAATCATCAGCACTGCTCTGAGGAGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 212

```
Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
Qy 213 AAACATTTGCTACTGATGTCGGCGATGTCATATAAATCTGAGAAAGCTGACTACTA 272
Db 61 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeu 80
Qy 273 AAGGCTATTCAAGAACAAATGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 332
Db 81 LysAlaileGlnGluLeuileAlaAsnValHisSerAsnAspGlyTyrPheGluVal 100
Qy 333 ATTGATTTTGCAGCGATCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Qy 393 AAGATGTTTGGTACCTTTCGGACCCCAACCTGTCACAAATTTTTCGCTAAAGCGACAT 452
Db 121 ArgAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
Qy 453 GTGCGGTTAGCCATATATAAGAAACCAATACAAACCAACGAAATCTGTTGATGTG 512
Db 141 ValArgValArgProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 160
Qy 513 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572
Db 161 AsnTyrGluValSerPheValSerGluThrGlyAsnLeuAspPheThrProSerLeuLys 180
Qy 573 GATACCTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCATCAAGAAATACTA 632
Db 181 GluArgTyrHisLeuThrThrLeuAlaValAlGlyAspSerLeuSerGlnGluLeuAla 200
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 692
Db 201 AlaIleAlaGlnPheIleLeuSerLysGluHisProAspTyrIleIleThrLysArgAsp 220
Qy 693 TCCTCAATCGTCATGACAAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 752
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 753 TTTACTTACCGTGTAAATATCGGACGATTTATAGCTTATAGATCAATAAAAATCTGCTG 812
Db 241 PheThrTyrHisIleLysAspArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 260
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 872
Db 261 ValGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
Qy 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTCAAACTGTTCACCAATCAATACGTT 932
Db 281 GluGluProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
Qy 933 GATGTCGATACCAACGAATTGCTAAAGTGAAGAGCTCTTAACAGCTAGCGAACCTAAC 992
Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
Qy 993 TTGACTTCAGAGATTATACCATCTCTGATAGAGTAAAGTAACTCTTCAACATCTC 1052
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
Qy 1053 GATGCTTTGTTATGATGACTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC 1112
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsnAsp 360
Qy 1113 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAC 1172
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
Qy 1173 CATTTAGCT 1181
Db 381 HisLeuAla 383
RESULT 6
S04168
streptokinase A precursor - Streptococcus pyogenes (strain NZ131)
```

```
C:Species: Streptococcus pyogenes
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: S04168
R:Huang, T.T.; Malke, H.; Ferretti, J.J.
Mol. Microbiol. 3, 197-205, 1989
A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri
A:Reference number: S04168; MUID:89343623; PMID:2668686
A:Accession: S04168
A:Molecule type: DNA
A:Residues: 1-440 <HUA>
A:Cross-references: UNIPROT:Q57391; UNIPARC:UPI00000BD04A; EMBL:X51517; NID:g47437; PID
C:Genetics:
C:Gene: ska
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-440/Product: streptokinase A #status predicted <MAT>
```

Alignment Scores:

Pred. No.:	5.73e-106	Length:	440
Score:	1670.00	Matches:	324
Percent Similarity:	90.34%	Conservative:	22
Best Local Similarity:	84.60%	Mismatches:	37
Query Match:	61.49%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x S04168 (1-440)

```
Qy 33 ATAGCTGGTCTGTAATGCTGCTAGATCGTCCTTCTGTAATAACAGCCAAATCGTGTGT 92
Db 27 IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
Qy 93 AGCGTCTGCTGCTACTGTTGAGGCGCAATCAAGACATAGTCTTAATATTTTGAATC 152
Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnIleSerLeuLysPhePheGluIle 66
Qy 153 GATCTTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAGTCCAAATCA 212
Db 67 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 86
Qy 213 AAACATTTGCTACTGATGTCGGCGATGTCATATAAATCTTGAGAAGCTGACTACTA 272
Db 87 LysProPheAlaThrAspLysGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 106
Qy 273 AAGGCTATTCAAGAACAAATGATCGCTTAAAGTCCACAGTAAACGACGACTCTTTGAGGTC 332
Db 107 LysAlaileGlnGluGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Qy 333 ATTGATTTTGCAGCGATGCAACATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
Qy 393 AAGATGTTTGGTACCTTTCGGACCCCAACCTGTCACAAATTTTTCGCTAAAGCGACAT 452
Db 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Qy 453 GTGCGGCTTAGACATATAAGAAACCAATACAAACCAACGAGAAATCTGTTGATGTG 512
Db 167 ValArgValLysProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 186
Qy 513 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGCTCAAA 572
Db 187 AsnTyrGluValSerPheValSerGluThrGlyAspLeuAspPheThrProLeuLeuArg 206
Qy 573 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATACTA 632
Db 207 AsnGlnTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerGlnGluLeuAla 226
Qy 633 GCTCAAGCACAAAGCATTTTAAACCAACCAACCCAGGCTATACGATTTATGAAACGTGAC 692
Db 227 AlaIleAlaGlnPheIleLeuSerLysLysHisProAspTyrIleIleThrLysArgAsp 246
Qy 693 TCCTCAATCGTCTCATGACCAATGACATTTTCGTCGATCGATTTTACCAATGGATCAAGAG 752
```

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 753 TTTACTTACCGTGTAAATAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 812
 Db 267 PheThrThrHisIleLysAspArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 286
 QY 813 AATGAAGAAATAACAACACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
 Db 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 873 GAAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTTCAACATCAAAATACGTT 932
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
 QY 933 GATGTCGATACCAACGAATTGCTAAAGTGAAGCTCTTAAACAGCTAGCGAAGCTAAC 992
 Db 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 993 TTAGACTTTCAGAGATTATACGATCTCTGATAAGGCTAAACTACTTACCAACATCTC 1052
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1053 GATGTTTTGGTATTATGACTATACCTTAACTGAAAGTAGAGGATTAATCAAGATGAC 1112
 Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1113 ACCAAGCTATCATACCGTTTATATGGCAGACGACCGAAGGAGAGATGCTAGCTAC 1172
 Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1173 CATTTAGCT 1181
 Db 407 HisLeuAla 409

RESULT 7

S53334
 streptokinase - Streptococcus sp.
 C;Species: Streptococcus sp.
 C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 C;Accession: S53334
 R;Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
 Biochem. J. 304, 235-241, 1994
 A;Title: Function of streptokinase fragments in plasminogen activation.
 A;Reference number: S53334; MUID:95091634; PMID:7998939
 A;Accession: S53334
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-197 <SH1>
 A;Cross-references: UNIPROT:Q7M115; UNIPARC:UPI000017AC2C

Alignment Scores:
 Pred. No.: 8.69e-38 Length: 197
 Score: 667.00 Matches: 167
 Percent Similarity: 44.71% Conservative: 2
 Best Local Similarity: 44.18% Mismatches: 1
 Query Match: 24.56% Indels: 208
 DB: 2 Gaps: 9

US-09-940-235-9 (1-1541) x S53334 (1-197)

QY 33 ATAGCTGTCTGAATGGCTACTAGATCGTCTCTGTAAATAACAGCCAAATGGTTGT 92
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsn----- 14
 QY 93 AGCGTTGTGCTACTGTGTTGAGGGGAGCAATCAAGCAATAGTCTTAAATTTTGGAAATC 152
 Db 14 ----- 14
 QY 153 GATCTTAACATCAAGCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 212
 Db 15 -----Ser 15
 QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTACTA 272

Db 16 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 35
 QY 273 AAGGCTATTCAAGAACAAATTGATCGTACGTCCACAGTAAACGACGACTACTTTGAGTTC 332
 Db 36 LysAlaIleGlnTrpGlnIleLeu----- 43
 QY 333 ATTGATTTTGAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
 Db 44 -----AsnGlyLysValTyrPheAlaAsp 51
 QY 393 AAAGATGGTTCGGTAAACCTTTCGCGACCAACCTGTCCAAGANTTTTTCGTAAGCGGCAT 452
 Db 52 LysAspGlySerValThr----- 57
 QY 453 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGGAAATCTGTTGATGTG 512
 Db 58 -----GluLysProIleGlnAsnGlnAlaLysSerValAspVal 70
 QY 513 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAA 572
 Db 71 Glu----- 71
 QY 573 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 632
 Db 71 ----- 71
 QY 633 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGATGAACGTC 692
 Db 72 -----AsnHisProGlyTyrThrIleTyrGluArgAsp 82
 QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGTCAGCATTTTACCAATGGATCAAGAG 752
 Db 83 SerSerIleVal-----ThrIleLeuProMetAspGlnGlu 94
 QY 753 TTTACTTACCGTGTATAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCTG 812
 Db 95 PheThrTyr-----SerGlyLeu 100
 QY 813 AATGAAGAAATAACAACACTGACTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
 Db 101 AsnGluGluIleAsnAsnThrAspLeuIleSer----- 111
 QY 873 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCACCAATAGCTT 932
 Db 112 -----TyrVal 113
 QY 933 GATGTCGATACCAACGAATTGTATAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 992
 Db 114 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 133
 QY 993 TTAGACTTCAGAGATTATACGATCTCTGTCGTAAGCTAAACTACTCTACACAATCTC 1052
 Db 134 -----AspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 149
 QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATTAATCAGATGAC 1112
 Db 150 AspAlaPheGlyIleMet----- 155
 QY 1113 ACCAAGCTATCATACCGTTTATATGGCAACGACCGCCGAGGAGAGAAATGCT 1166
 Db 156 -----IleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAla 170

RESULT 8

FNHU
 fibronectin precursor [validated] - human
 N;Alternate names: fibronectin splice form ED-A
 C;Species: Homo sapiens (man)
 C;Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2004
 C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A232
 R;Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
 A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656
 A;Accession: A26460
 A;Molecule type: DNA
 A;Residues: 1-49 <DEA>
 A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; UNIPARC:UPI000016A926; GB:M15801; NID:1494
 R;Oldberg, A.; Ruoslahti, E.
 J. Biol. Chem. 261, 2113-2116, 1986
 A;Title: Evolution of the fibronectin gene.
 A;Reference number: A26284; MUID:86111901; PMID:3003095
 A;Accession: A26284
 A;Molecule type: DNA
 A;Residues: 1447-1540 <OLD>
 A;Cross-references: UNIPARC:UPI0000112E37; GB:M12549; NID:g182688
 A;Note: the authors translated the codon TTC for residue 1494 as Glu
 R;Paolella, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
 Nucleic Acids Res. 16, 3545-3557, 1988
 A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-E
 A;Reference number: S00848; MUID:88233940; PMID:3375063
 A;Accession: S03917
 A;Molecule type: DNA
 A;Residues: 1594-1767, 'V', 1769-1783 <PAO>
 A;Cross-references: UNIPARC:UPI000017432B; EMBL:X07718; NID:g31402
 A;Note: the authors translated the codon AAC for residue 1631 as Asp
 R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 FEBS Lett. 207, 287-291, 1986
 A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.
 A;Reference number: A24854; MUID:87030929; PMID:3770201
 A;Accession: A24854
 A;Molecule type: DNA
 A;Residues: 1992-2147 <VIB>
 A;Cross-references: UNIPARC:UPI000017432C; GB:X04530; NID:g31436
 R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.
 FEBS Lett. 207, 145-148, 1986
 A;Title: Human fibronectin is synthesized as a pre-polypeptide.
 A;Reference number: A24476; MUID:87030890; PMID:3770189
 A;Accession: A24476
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-14, 'Q', 16-38 <GUT>
 A;Cross-references: UNIPARC:UPI000017432D
 R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A;Title: Primary structure of human fibronectin: differential splicing may generate at l
 A;Reference number: A91008; MUID:85284965; PMID:2992939
 A;Accession: A91008
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
 A;Cross-references: UNIPARC:UPI000017432E; UNIPARC:UPI000017432F; GB:X02761
 R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
 A;Reference number: A93529; MUID:84272258; PMID:6462919
 A;Accession: A93529
 A;Molecule type: mRNA
 A;Residues: 973-2080; 2112-2386 <KO2>
 A;Cross-references: UNIPARC:UPI000017432F; UNIPARC:UPI0000174330; GB:X00739
 R;Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
 A;Reference number: A21011; MUID:83290929; PMID:6688418
 A;Accession: A21011
 A;Molecule type: mRNA
 A;Residues: 1434-1537 <OL2>
 A;Cross-references: UNIPARC:UPI0000174331; GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:1
 R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2698-2704, 1985
 A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
 A;Reference number: A90495; MUID:85280409; PMID:2992573
 A;Accession: A90495
 A;Molecule type: mRNA
 A;Residues: 1594-2386 <BER>
 A;Cross-references: UNIPARC:UPI000016A928; GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:1

R;Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.
 FEBS Lett. 186, 31-34, 1985
 A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
 A;Reference number: A22245; MUID:85231203; PMID:2989004
 A;Accession: A22245
 A;Molecule type: mRNA
 A;Residues: 1948-2067 <UME>
 A;Cross-references: UNIPARC:UPI000000046A; GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:1
 A;Accession: B22245
 A;Molecule type: mRNA
 A;Residues: 1975-1991, 2017-2039 <UM2>
 A;Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
 R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
 Biochemistry 25, 4936-4941, 1986
 A;Title: Human liver fibronectin complementary DNAs: identification of two different me
 A;Reference number: 152394; MUID:87026578; PMID:3021206
 A;Accession: 165273
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
 A;Cross-references: UNIPARC:UPI000006E04C; GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:1
 R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
 A;Reference number: A21165; MUID:83221567; PMID:6304699
 A;Accession: A21165
 A;Molecule type: mRNA
 A;Residues: 2291-2386 <KO3>
 A;Cross-references: UNIPARC:UPI0000174334; GB:X00799; NID:g182681; PIDN:AAA52460.1; PID:1
 R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
 J. Biol. Chem. 258, 12670-12674, 1983
 A;Title: Primary structure of human plasma fibronectin.
 A;Reference number: A92398; MUID:84032463; PMID:6630202
 A;Accession: A92398
 A;Molecule type: protein
 A;Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
 A;Cross-references: UNIPARC:UPI0000174335
 R;Garcia-Pardo, A.; Gold, L.I.
 Arch. Biochem. Biophys. 304, 181-188, 1993
 A;Title: Further characterization of the binding of fibronectin to gelatin reveals the p
 A;Reference number: S34791; MUID:93312001; PMID:8323285
 A;Accession: S34791
 A;Molecule type: protein
 A;Residues: 291-300, 551-560 <GAR2>
 A;Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
 R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
 Thromb. Res. 43, 469-477, 1986
 A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
 A;Reference number: A60904; MUID:87019725; PMID:3532418
 A;Accession: A60904
 A;Molecule type: protein
 A;Residues: 293-301 <GRI>
 A;Cross-references: UNIPARC:UPI0000174338
 R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
 J. Biol. Chem. 260, 12136-12141, 1985
 A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p
 A;Reference number: A23901; MUID:86008277; PMID:3900070
 A;Accession: A23901
 A;Molecule type: protein
 A;Residues: 616-677, 'Q', 679-703, 'PT' <CAL>
 A;Cross-references: UNIPARC:UPI0000174339
 R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
 J. Biol. Chem. 257, 9593-9597, 1982
 A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
 A;Reference number: A92386; MUID:82265604; PMID:7050098
 A;Accession: A92386
 A;Molecule type: protein
 A;Residues: 1441-1548 <FIE>
 A;Cross-references: UNIPARC:UPI0000141CD5
 A;Note: Residues 1524-1527 are responsible for the cell-binding activity
 R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
 Biochem. J. 241, 923-928, 1987
 A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom

A;Reference number: A32517; MUID:87241275; PMID:3593230

A;Accession: A32517

A;Molecule type: protein

A;Residues: 1589-1630, 'T', 1722-2058 <GAR3>

A;Cross-references: UNIPARC:UPI000017433A

R;Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand

Biochem. J. 274, 731-738, 1991

A;Title: Human plasma fibronectin. Demonstration of structural differences between the A

A;Reference number: S14357; MUID:91190085; PMID:2012601

A;Accession: S14357

A;Molecule type: protein

A;Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>

A;Cross-references: UNIPARC:UPI000017433B

R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.

J. Biol. Chem. 260, 10320-10325, 1985

A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal

A;Reference number: A23891; MUID:85261459; PMID:4019516

A;Accession: A23891

A;Molecule type: protein

A;Residues: 2071-2080, 2112-2356 <GAR4>

A;Cross-references: UNIPARC:UPI000017433C; UNIPARC:UPI000017433D

C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis

C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,

ation, and transformation.

C;Genetics:

A;Gene: GDB:FNI

A;Cross-references: GDB:119135; OMIM:135600

A;Map position: 2q34-q34

A;Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1

C;Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology;

C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-31/Domain: propeptide #status predicted <PRO>

F;32-2386/Product: fibronectin #status experimental <MAT>

F;52-272/Domain: fibrin and heparin binding <FHB>

F;52-87/Domain: fibronectin type I repeat homology <1F1>

F;97-135/Domain: fibronectin type I repeat homology <1F2>

F;141-179/Domain: fibronectin type I repeat homology <1F3>

F;186-225/Domain: fibronectin type I repeat homology <1F4>

F;231-270/Domain: fibronectin type I repeat homology <1F5>

F;308-342/Domain: collagen binding <CBR>

F;360-401/Domain: fibronectin type II repeat homology <2F1>

F;420-461/Domain: fibronectin type II repeat homology <2F2>

F;470-508/Domain: fibronectin type I repeat homology <1F7>

F;518-555/Domain: fibronectin type I repeat homology <1F8>

F;561-599/Domain: fibronectin type I repeat homology <1F9>

Alignment Scores:

Pred. No.:	2,618-35	Length:	2386
Score:	632.00	Matches:	117
Percent Similarity:	87.59%	Conservative:	3
Best Local Similarity:	85.40%	Mismatches:	11
Query Match:	23.27%	Indels:	6
DB:	1	Gaps:	3

US-09-940-235-9 (1-1541) x FNHU (1-2386)

QY 1143 AAGGACCCGAGGAGAGAGATCTAGCTACCACTTTA-----GCTGCTGTGGC 1190

Db 155 ArgArgProHis---GluThrGlyTyrMetLeuGlyCysValCysLeuGlyAsnGly 173

QY 1191 CAGGCGCAA---CAGATTGTACCATAGCTGAGAGAGTGTGTGATCATGCTGCTGGACT 1247

Db 174 LysGlyGluThrTrpCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193

QY 1248 TCCTATGTCGCGGAACCTGGAGAGCCCTACCAAGCTGGATGATGCTAGATTCT 1307

Db 194 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetValAspCys 213

QY 1308 ACTTGCCCTGGAGAGGAGCGGAGCGCATCTTGACATCTTAGAAATAGATCAACGAT 1367

Db 214 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 233

QY 1368 CAGGACACAGGACATCCTATAGATTGGAGACACCTGGAGCAAGAGGATAATCGAGGA 1427

Db 234 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 253

QY 1428 AACCTGCTCCAGTGCATCTGCACAGCAACCGCGGAGGAGAGTGGAAGTGTGAGAGCAC 1487

Db 254 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273

QY 1488 ACCTCTGTGCAGACCATCGAGGATCTGGCCCTTACCCGATGTCGT 1538

Db 274 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 290

RESULT 9

S77671

streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)

C;Species: Streptococcus pyogenes

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999

C;Accession: S77671; S77672

R;Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,

Mol. Microbiol. 16, 509-519, 1995

A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A;Reference number: S77671; MUID:96037795; PMID:7565111

A;Accession: S77671

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAP>

A;Cross-references: UNIPARC:UPI0000000562; EMBL:U25853; NID:g818908; PIDN:AAA85729.1; P

A;Experimental source: strain ET1/M1

A;Note: allele 2

A;Accession: S77672

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAX>

A;Cross-references: UNIPARC:UPI0000000562; EMBL:U25854; NID:g818910; PIDN:AAA85730.1; P

A;Experimental source: strain E2/M3

A;Note: allele 3

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	1,318-34	Length:	128
Score:	620.00	Matches:	118
Percent Similarity:	95.31%	Conservative:	4
Best Local Similarity:	92.19%	Mismatches:	6
Query Match:	22.83%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-9 (1-1541) x S77671 (1-128)

QY 453 GTGCGGTTAGACATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512

Db 1 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 20

QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTATGACGATTCAGACAGGTCTCAA 572

Db 21 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 40

QY 573 GATACCTAAGCTATTGAACACCTAGCTATCGGTGACACCATCATCTCAAGAACTACTA 632

Db 41 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 60

QY 633 GCTCAAGCAACAAAGCATTTTAAACAAAAACCCACAGGCTATACGATTTATGAACGTGAC 692

Db 61 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 80

QY 693 TCCTCAATCGTCACTCATGACAAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100

QY 753 TTACTTACCGTGTAAATAATCGGAAACAGCTTAGGATCAATAAAAAATCTGGTCTG 812

```

Db      101 PheThrTyHisValysAsnArgGluGlnAlaTyRGluIleAsnProIysThrGlyIle 120
QY      813 AATGAAGAAATAAACCAACTGAC 836
Db      121 LysGluLysThrAsnAsnThrAsp 128

RESULT 10
FNBO
fibronectin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A26452; B21165; A23292
R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.
Eur. J. Biochem. 161, 441-453, 1986
A:Title: Complete primary structure of bovine plasma fibronectin.
A:Reference number: A26452; MUID:87054047; PMID:3780752
A:Accession: A26452
A:Molecule type: protein
A:Residues: 1-2265 <SKO>
A:Cross-references: UNIPROT:P07589; UNIPARC:UPI000012A7BE
R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A:Reference number: A21165; MUID:83221567; PMID:6304699
A:Accession: B21165
A:Molecule type: mRNA
A:Residues: 2170-2265 <KOR>
A:Cross-references: UNIPARC:UPI000016C30C; GB:X00800; NID:9163055; PIDN:AAA30521.2; PID:
R:Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottr
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal
A:Reference number: A23292; MUID:83117805; PMID:6218503
A:Accession: A23292
A:Molecule type: protein
A:Residues: 1-16, 'C', 18-20, 'S', 22-432, 447-463, 1367-1517, 1567-1673, 2062-2176, 'N', 2178-226
A:Cross-references: UNIPARC:UPI000017433E; UNIPARC:UPI000017433F; UNIPARC:UPI0000174340;
C:Comment: Cys-1201 and Cys-2015 have free sulphydryl groups.
C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b
C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib
aling, and maintenance of cell shape.
C:Comment: Plasma fibronectin is synthesized by hepatocytes.
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu
F:21-56/Domain: fibronectin type I repeat homology <1F1>
F:66-104/Domain: fibronectin type I repeat homology <1F2>
F:110-148/Domain: fibronectin type I repeat homology <1F3>
F:155-194/Domain: fibronectin type I repeat homology <1F4>
F:200-239/Domain: fibronectin type I repeat homology <1F5>
F:277-577/Domain: collagen binding <CBR>
F:277-311/Domain: fibronectin type I repeat homology <1F6>
F:329-370/Domain: fibronectin type II repeat homology <2F1>
F:389-430/Domain: fibronectin type II repeat homology <2F2>
F:439-477/Domain: fibronectin type I repeat homology <1F7>
F:487-524/Domain: fibronectin type I repeat homology <1F8>
F:530-568/Domain: fibronectin type I repeat homology <1F9>
F:578-661/Domain: fibronectin type III repeat homology <FN3A>
F:688-770/Domain: fibronectin type III repeat homology <FN3B>
F:779-860/Domain: fibronectin type III repeat homology <FN3C>
F:875-957/Domain: fibronectin type III repeat homology <FN3D>
F:965-1046/Domain: fibronectin type III repeat homology <FN3E>
F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F:1326-1404/Domain: fibronectin type III repeat homology <GN3I>
F:1410-1517/Domain: cell attachment <CAD>
F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F:1493-1495/Region: cell attachment (R-G-D) motif
F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F:1600-1870/Domain: heparin binding <HB2>
F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>
F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>

```

```

F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F:1970-1972/Region: cell attachment (R-G-D) motif
F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F:1985-2216/Domain: fibrin binding <FB2>
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyroliidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) interchain to Lys N6-amino of fibrin) #status experimental
F:21-47, 45-56, 66-94, 92-104, 110-138, 138-148, 155-184, 182-194, 200-229, 227-239, 277-304, 302-
7, 2155-2167, 2174-2200, 2198-2209/Disulfide bonds: #status predicted
F:399, 497, 511, 846, 976, 1213, 1987/Binding site: carbohydrate (Asn) (covalent) #status absent
F:1205, 1692/Binding site: carbohydrate (Asn) (covalent) #status absent
F:1943, 1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:2246/Disulfide bonds: interchain (to 2250) #status predicted
F:2250/Disulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Alignment Scores:
Pred. No.:      8,14e-34      Length:      2265
Score:          610.00      Matches:      113
Percent Similarity: 86.13%      Conservative: 5
Best Local Similarity: 22.46%      Mismatches: 13
Query Match:      1          Indels: 6
DB:              1          Gaps: 3

US-09-940-235-9 (1-1541) x FNBO (1-2265)
QY      1143 AAGCGACCCGAGGAGAGATGCTAGTACCATTTA-----GCTGGTGGTGGC 1190
Db      124 ArgArgProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly 142
QY      1191 CAGCGCGCAA---CAGATTGTACCATAGCTGAGAGAGTGTTTTGATCATGCTGCTGGGACT 1247
Db      143 LysGlyGluTrpThrCysLysProIleAlaGluLysCysPheAspGlnAlaAlaGlyThr 162
QY      1248 TCCTATGTGTCGGAGAAACCTGGGAAAGCCCTACCAAGCTCGATGATGGTAGATTGT 1307
Db      163 SerTyrValValGlyGluThrTrpGluLysProIleGlyTrpMetMetValAspCys 182
QY      1308 ACTTGTCTGGGAGAGGACGACGACGATCCTTGCACTTGTGCAATCTTCTAGAAATAGATGCAACGAT 1367
Db      183 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 202
QY      1368 CAGACACACAGGACATCTTAGAATTGGAGACACCTGGAGCAAGAGGATATATCGAGGA 1427
Db      203 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 222
QY      1428 AACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGAGAGTGGAAGTGTGTGAGAGGCAC 1487
Db      223 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
QY      1488 ACCTCTGTGACACACCATCGAGCGGATCTGGCCCTTCCACCGATGTTGTT 1538
Db      243 ThrSerLeuGlnThrThrSerAlaGlySerGlySerPheThrAspValArg 259

RESULT 11
S14428
fibronectin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
R:Hynes, R.O.
submitted to the EMBL Data Library, July 1989
A:Reference number: S14428
A:Accession: S14428
A:Molecule type: mRNA
A:Residues: 1-2477 <HYN>
A:Cross-references: UNIPROT:P04937; UNIPARC:UPI000012A7C6; EMBL:X15906; NID:956163; PIDN:
R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
EMBO J. 6, 2573-2580, 1987
A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
A:Reference number: S12455; MUID:88054951; PMID:2445560

```


A:Accession: S12455
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 609-1810,'T',1812-2283 <SCH>
A:Cross-references: UNIPARC:UPI0000177AEC; EMBL:X15906
R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
A:Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing
A:Reference number: A23219; MUID:84298097; PMID:6089177
A:Accession: A23219
A:Molecule type: DNA
A:Residues: 2052-2237 <TAM>
A:Cross-references: UNIPARC:UPI0000177AEB
R:Falkenberg, C.; Enghild, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstrøm, B.
Biochem. J. 301, 745-751, 1994
A:Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in the presence of a monoclonal antibody
A:Reference number: S46203; MUID:94330948; PMID:7519849
A:Accession: S46203
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1183-1192,'GLN',1268,'P',1270-1271,'D',1273,'CF',1276,'PY',1385-1399 <FAL>
A:Cross-references: UNIPARC:UPI0000177AEC; UNIPARC:UPI0000177AED; UNIPARC:UPI0000177AEE
R:Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.
EMBO J. 6, 2565-2572, 1987
A:Title: Organization of the fibronectin gene provides evidence for exon shuffling during evolution
A:Reference number: S00459; MUID:88054950; PMID:3119323
A:Accession: S00459
A:Molecule type: DNA
A:Residues: 1-139,2382-2477 <PAT>
A:Cross-references: UNIPARC:UPI000017095C; UNIPARC:UPI0000177AEP; EMBL:X05831
A:Note: the authors translated the codon CCT for residues 51 and 94 as Ala
R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.
Cell 35, 421-431, 1983
A:Title: Three different fibronectin mRNAs arise by alternative splicing within the coding region
A:Reference number: A27253; MUID:84082067; PMID:6317187
A:Accession: A27253
A:Molecule type: mRNA
A:Residues: 1586-1720,'T',1722,1813-2477 <SC2>
A:Cross-references: UNIPARC:UPI0000177AFO
R:Odermatt, E.; Tamkun, J.W.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
A:Title: Repeating modular structure of the fibronectin gene: Relationship to protein structure
A:Reference number: I59049; MUID:86016741; PMID:3863113
A:Accession: I59049
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1722-1810 <RES>
A:Cross-references: UNIPARC:UPI000000040C; GB:M11750; MUID:9204164; PIDN:AAA41170.1; PID: I59049
C:Genetics:
C:Introns: 51/1; 94/1; 2416/3; 2454/3
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology; fibronectin type III repeat homology; collagen binding; disulfide bond; duplication
C:Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; duplication
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-2477/Product: fibronectin type I repeat homology <1F1>
F:53-88/Domain: fibronectin type I repeat homology <1F2>
F:98-136/Domain: fibronectin type I repeat homology <1F3>
F:142-180/Domain: fibronectin type I repeat homology <1F4>
F:187-226/Domain: fibronectin type I repeat homology <1F5>
F:232-271/Domain: fibronectin type I repeat homology <1F6>
F:308-342/Domain: fibronectin type I repeat homology <1F7>
F:360-401/Domain: fibronectin type I repeat homology <1F8>
F:420-461/Domain: fibronectin type I repeat homology <1F9>
F:470-508/Domain: fibronectin type I repeat homology <1F10>
F:518-555/Domain: fibronectin type I repeat homology <1F11>
F:561-599/Domain: fibronectin type I repeat homology <1F12>
F:609-632/Domain: fibronectin type I repeat homology <1F13>
F:718-800/Domain: fibronectin type I repeat homology <1F14>
F:809-830/Domain: fibronectin type I repeat homology <1F15>
F:905-987/Domain: fibronectin type I repeat homology <1F16>
F:995-1076/Domain: fibronectin type I repeat homology <1F17>
F:1085-1164/Domain: fibronectin type I repeat homology <1F18>
F:1172-1257/Domain: fibronectin type I repeat homology <1F19>
F:1365-1348/Domain: fibronectin type I repeat homology <1F20>

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 3.86e-31 Length: 128
Score: 569.00 Matches: 110
Percent Similarity: 91.41% Conservative: 7
Best Local Similarity: 85.94% Mismatches: 11
Query Match: 20.95% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S77680 (1-128)

QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 512

Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20

QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572

Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 40

QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 632

Db 41 AsnThrLysLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 60

QY 633 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692

Db 61 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 80

QY 693 TCCTCAATCGTCACATGACATGATTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100

QY 753 TTTACTTACCGTGTAAAAATCGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Db 101 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyLeuAsnLysSerGlyGln 120

QY 813 AATGAAGAAATAAACACACTGAC 836

Db 121 GluGluLysThrAsnAsnThrAsp 128

RESULT 13

S77688

streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)

C;Species: Streptococcus pyogenes

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C;Accession: S77688

R;Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,

Mol. Microbiol. 16, 509-519, 1995

A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A;Reference number: S77671; MUID:96037795; PMID:7565111

A;Accession: S77688

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAP>

A;Cross-references: UNIPROT:Q54695; UNIPARC:UPI00000BD8DC; EMBL:U25870; NID:9818942; PII

A;Experimental source: strain ET76/M72

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 2.53e-30 Length: 128
Score: 557.00 Matches: 107
Percent Similarity: 89.84% Conservative: 8
Best Local Similarity: 83.59% Mismatches: 13
Query Match: 20.51% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S77688 (1-128)

QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 512

Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20

QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572

Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 40

QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 632

Db 41 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 60

QY 633 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692

Db 61 AlaGlnAlaGlnSerIleLeuLysGluSerHisProAspTyrThrIleTyrGluArgAsp 80

QY 693 TCCTCAATCGTCACATGACATGATTTCCGTACGATTTTACCAATGGATCAAGAG 752

Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100

QY 753 TTTACTTACCGTGTAAAAATCGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812

Db 101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspVal 120

QY 813 AATGAAGAAATAAACACACTGAC 836

Db 121 LysGluLysThrAsnAsnThrAsp 128

RESULT 14

S77679

streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)

C;Species: Streptococcus pyogenes

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C;Accession: S77679

R;Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,

Mol. Microbiol. 16, 509-519, 1995

A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A;Reference number: S77671; MUID:96037795; PMID:7565111

A;Accession: S77679

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAP>

A;Cross-references: UNIPROT:Q54686; UNIPARC:UPI00000BD8D6F; EMBL:U25861; NID:9818924; P

A;Experimental source: strain ET50/M43

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 5.53e-30 Length: 128
Score: 552.00 Matches: 106
Percent Similarity: 89.06% Conservative: 8
Best Local Similarity: 82.81% Mismatches: 14
Query Match: 20.32% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x S77679 (1-128)

QY 453 GTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 512

Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20

QY 513 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572

Db 21 ArgTyrAlaValGlnPheThrProLeuAsnProAspAspPheThrProValLeuLys 40

QY 573 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 632

Db 41 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 60

QY 633 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692

```

Db      61 AlaGlnAlaGlnSerIleLeuIleGluSerHisProAsnTyrThrIleHisGluArgAsp 80
      693 TCCTCAATCGTCACATGACAAATGATTTTCCTGACGATTTTACCAATGGATCAAGAG 752
      81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100
      753 TTTACTTACCGTGTAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 812
      101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 120
      813 AATGAAGAAATAAACCACTGAC 836
      121 LysGluLysThrAsnAsnThrAsp 128

RESULT 15
A43908
Fibronectin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
C:Accession: A43908
R:DeSimone, D.W.; Norton, P.A.; Hynes, R.O.
Dev. Biol. 149, 357-369, 1992
A:Title: Identification and characterization of alternatively spliced fibronectin mRNAs
A:Reference number: A43908; MUID:92111942; PMID:1730390
A:Accession: A43908
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2481 <DB>
A:Cross-references: UNIPARC:UPI0000177AE7; GB:M77820
C:Note: sequence extracted from NCBI backbone (NCBIP:77473)
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C:Keywords: duplication; extracellular matrix; glycoprotein; heterodimer
F:55-90/Domain: fibronectin type I repeat homology <1F1>
F:100-138/Domain: fibronectin type I repeat homology <1F2>
F:144-182/Domain: fibronectin type I repeat homology <1F3>
F:189-228/Domain: fibronectin type I repeat homology <1F4>
F:234-273/Domain: fibronectin type I repeat homology <1F5>
F:309-343/Domain: fibronectin type I repeat homology <1F6>
F:361-402/Domain: fibronectin type II repeat homology <2F1>
F:421-462/Domain: fibronectin type II repeat homology <2F2>
F:471-509/Domain: fibronectin type I repeat homology <1F7>
F:519-556/Domain: fibronectin type I repeat homology <1F8>
F:562-600/Domain: fibronectin type I repeat homology <1F9>
F:610-693/Domain: fibronectin type III repeat homology <FN3A>
F:719-801/Domain: fibronectin type III repeat homology <FN3B>
F:810-891/Domain: fibronectin type III repeat homology <FN3C>
F:906-988/Domain: fibronectin type III repeat homology <FN3D>
F:996-1077/Domain: fibronectin type III repeat homology <FN3E>
F:1086-1165/Domain: fibronectin type III repeat homology <FN3F>
F:1173-1258/Domain: fibronectin type III repeat homology <FN3G>
F:1266-1349/Domain: fibronectin type III repeat homology <FN3H>
F:1357-1440/Domain: fibronectin type III repeat homology <FN3I>
F:1448-1530/Domain: fibronectin type III repeat homology <FN3J>
F:1538-1620/Domain: fibronectin type III repeat homology <FN3K>
F:1615-1617/Region: cell attachment (R-G-D) motif
F:1632-1714/Domain: fibronectin type III repeat homology <FN3L>
F:1722-1804/Domain: fibronectin type III repeat homology <FN3M>
F:1812-1894/Domain: fibronectin type III repeat homology <FN3N>
F:1904-1985/Domain: fibronectin type III repeat homology <FN3O>
F:1993-2075/Domain: fibronectin type III repeat homology <FN3P>
F:2197-2277/Domain: fibronectin type III repeat homology <FN3Q>
F:2301-2340/Domain: fibronectin type I repeat homology <1F10>
F:2346-2383/Domain: fibronectin type I repeat homology <1F11>
F:2390-2425/Domain: fibronectin type I repeat homology <1F12>
F:55-81,79-90,100-128,126-138,144-172,170-182,189-218,216-228,234-263,261-273,309-336,33
2373,2371-2383,2390-2416,2414-2425/Disulfide bonds: #status predicted
F:2459/Disulfide bonds: interchain (to 2463) #status predicted
F:2463/Disulfide bonds: interchain (to 2459) #status predicted

```

Alignment Scores:

Pred. No.:	2,53e-29	Length:	2481
Score:	544.00	Matches:	95

Percent Similarity: 81.75%
 Best Local Similarity: 69.34%
 Query Match: 20.03%
 DB: 2
 Conservative: 17
 Mismatches: 19
 Indels: 6
 Gaps: 3

US-09-940-235-9 (1-1541) x A43908 (1-2481)

```

QY      1143 AAGCGACCCGAGGAGGAGATGCTAGCTACCATTTA-----GCTGGTGGTGGC 1190
      158 AGArgprOHis---GluThrGlyGlyTyrMeLeuGluCysValCysLeuGlyAsnGly 176
      1191 CAGCGCAA---CAGATTGTACCCATAGCTGAGAGTGTGTTTGTATCTGCTGGGACT 1247
      177 LysGlyGluTrpThrCysLysProValAlaGluArgCysTyrAspAsnThrAlaGlyThr 196
      1248 TCCTATGCTGTCGGAGAAACGTCGGAGAGCCCTACCAAGCTGGATGGTAGATTGT 1307
      197 SerTyrValValGlyGlnThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 216
      1308 ACTTGCTGCGGAGAGCGGACGCATCTGACATCTCTAGAAATAGATGCAACGAT 1367
      217 ThrCysLeuGlyGluGlyAsnGlyArgIleThrCysSerSerLysAsnArgCysAsnAsp 236
      1368 CAGGACACAAGGACATCCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCCAGGA 1427
      237 GlnAspThrLysThrSerTyrArgIleGlyAspThrTrpSerLysThrAspThrArgGly 256
      1428 AACCTGCTCCAGTGCATCTGCAAGCAACCGCCGAGGAGAGTGGAGAGTGTGAGAGCAC 1487
      257 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 276
      1488 ACCTCTGTGTCAGACACATCGAGCGCATCTGGCCCTTACCCGATGTTCT 1538
      277 SerSerAlaGlnAlaThrGlyThrGlySerAsnProIleThrAsnIleGln 293

```

Search completed: January 28, 2006, 02:37:22
 Job time : 38.6156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 166.071 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-9
Perfect score: 2716
Sequence: 1 ttgttgaacttaagaagg.....ccttcaccgatgttcgtag 1541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0940235/runat_27012006_144218_27563/app_query.fasta_1.7708
-DB=UniProt -QFWT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0940235 @Cn_1_1_1359 @runat_27012006_144218_27563 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984	73.0	440	1 STRP_STRPQ	P00779 streptococc
2	1955	72.0	414	2 Q53284_STRPQ	Q53284 streptococc
3	1943	71.5	436	2 Q6UK57_STRPQ	Q6UK57 streptococc
4	1942	71.5	440	1 STRP_STRS1	P10519 streptococc
5	1814	66.8	440	2 Q7X0Y1_STRPY	Q7X0Y1 streptococc
6	1802	66.3	440	2 Q8X5R8_STRP3	Q8X5R8 streptococc
7	1802	66.3	440	1 STRP_STRPY	P10520 streptococc
8	1794	66.1	440	2 Q532X6_STRPY	Q532X6 streptococc
9	1794	66.1	440	2 Q7X0Y2_STRPY	Q7X0Y2 streptococc
10	1787	65.8	440	2 Q7X0Y7_STRPY	Q7X0Y7 streptococc
11	1758	64.7	440	2 Q5X9T6_STRP6	Q5X9T6 streptococc
12	1758	64.7	440	2 Q8NZA6_STRP8	Q8NZA6 streptococc
13	1755	64.6	440	2 Q7X0Y3_STRPY	Q7X0Y3 streptococc
14	1723	63.4	440	2 Q7X0Y0_STRPY	Q7X0Y0 streptococc
15	1715	63.1	440	2 Q7X0Y5_STRPY	Q7X0Y5 streptococc
16	1715	63.1	440	2 Q7X0Y5_STRPY	Q7X0Y5 streptococc

17	1708	62.9	440	2	Q7X0X8_STRPY	Q7X0X8 streptococc
18	1686	62.1	440	2	Q7X0X9_STRPY	Q7X0X9 streptococc
19	1677	61.7	440	2	Q7X0X7_STRPY	Q7X0X7 streptococc
20	1670	61.5	440	2	Q57391_STRPY	Q57391 streptococc
21	1657	61.0	440	2	Q7X0Y6_STRPY	Q7X0Y6 streptococc
22	1652	60.8	440	1	STRQ_STRPY	P96471 streptococc
23	1624	59.8	432	2	Q7X0Y4_STRPY	Q7X0Y4 streptococc
24	728	26.8	141	2	Q7X0X2_STRPQ	Q7X0X2 streptococc
25	727	26.8	141	2	Q7X0X3_STRPQ	Q7X0X3 streptococc
26	724	26.7	141	2	Q7X0W1_STRPQ	Q7X0W1 streptococc
27	723	26.6	141	2	Q7WS87_STRPQ	Q7WS87 streptococc
28	717	26.4	141	2	Q7X0W3_STRPQ	Q7X0W3 streptococc
29	717	26.4	141	2	Q7X0X4_STRPQ	Q7X0X4 streptococc
30	715	26.3	141	2	Q7X0W5_STRPQ	Q7X0W5 streptococc
31	714	26.3	141	2	Q7X0X5_STRPQ	Q7X0X5 streptococc
32	712	26.2	141	2	Q7X0X6_STRPQ	Q7X0X6 streptococc
33	711	26.2	141	2	Q7X0W2_STRPQ	Q7X0W2 streptococc
34	711	26.2	141	2	Q7X0W9_STRPQ	Q7X0W9 streptococc
35	710	26.1	141	2	Q7X0W6_STRPQ	Q7X0W6 streptococc
36	707	26.0	141	2	Q7X0W0_STRPQ	Q7X0W0 streptococc
37	705	26.0	141	2	Q7X0V5_STRPY	Q7X0V5 streptococc
38	704	25.9	141	2	Q7X0X0_STRPQ	Q7X0X0 streptococc
39	701	25.8	141	2	Q7X0W7_STRPQ	Q7X0W7 streptococc
40	697	25.7	141	2	Q7X0W4_STRPQ	Q7X0W4 streptococc
41	696	25.6	141	2	Q7X0W8_STRPQ	Q7X0W8 streptococc
42	693	25.5	141	2	Q7X0X1_STRPQ	Q7X0X1 streptococc
43	690	25.4	141	2	Q7X0R2_STRPY	Q7X0R2 streptococc
44	686	25.3	141	2	Q7X0S4_STRPY	Q7X0S4 streptococc
45	684	25.2	141	2	Q7X0S0_STRPY	Q7X0S0 streptococc

ALIGNMENTS

RESULT 1

ID	STRP_STRPQ	STANDARD;	PRT;	440 AA.
AC	P00779;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Streptokinase C precursor.			
GN	Name=skc;			
OS	Streptococcus equisimilis.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=119602;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=H46A;			
EX	MEDLINE=85232082; PubMed=2989113; DOI=10.1016/0378-1119(85)90145-3;			
RA	Malke H., Roe B.A., Ferretti J.J.;			
RT	"Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A.";			
RL	Gene 34:357-362(1985).			
RN	[2]			
RP	PROTEIN SEQUENCE OF 27-440.			
RX	MEDLINE=83127125; PubMed=6760891;			
RA	Jackson K.W., Tang J.;			
RT	"Complete amino acid sequence of streptokinase and its homology with serine proteases.";			
RL	Biochemistry 21:6620-6625(1982).			
CC	-/- FUNCTION: This protein is not a protease, but it activates plasminogen by complexing with it. As a potential virulence factor, it is thought to prevent the formation of effective fibrin barriers around the site of infection, thereby contributing to the invasiveness of the cells.			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
CC	-----			

DR EMBL; K02986; AAA26974.1; -; Genomic DNA.
 DR EMBL; X72832; CAA51351.1; -; Genomic DNA.
 DR PIR; A00967; BZSO.
 DR PIR; A22801; A22801.
 DR PDB; 1BML; X-ray; C/D=38-399.
 DR PDB; 1L4D; X-ray; B=40-173.
 DR PDB; 1L4Z; X-ray; B=27-173.
 DR PDB; 1LQO; X-ray; A/B/C/D=177-314.
 DR SMC; P00779; 38-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 DR 3D-structure; Direct protein sequencing; Plasminogen activation;
 KW Signal; Virulence.
 FT SIGNAL 1 26
 FT CHAIN 27 440 Streptokinase C.
 FT VARIAT 195 195 L -> D.
 FT VARIAT 207 207 D -> L.
 FT CONFLICT 298 300 EKY -> LEYK (in Ref. 2).
 FT CONFLICT 438 438 N -> D (in Ref. 2).
 FT STRAND 180 180
 FT STRAND 184 194
 FT TURN 199 200
 FT TURN 203 204
 FT STRAND 205 205
 FT STRAND 209 214
 FT TURN 216 217
 FT STRAND 219 221
 FT HELIX 222 236
 FT TURN 238 239
 FT STRAND 240 252
 FT TURN 253 254
 FT STRAND 259 260
 FT STRAND 268 270
 FT STRAND 275 275
 FT STRAND 278 280
 FT TURN 282 284
 FT STRAND 287 289
 FT STRAND 292 304
 FT TURN 305 306
 SQ SEQUENCE 440 AA; 50140 MW; 8PCLF22648ACC77A CRC64;
 Alignment Scores:
 Pred. No.: 2.21e-133 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.05% Indels: 0
 DB: 1 Gaps: 0
 US-09-940-235-9 (1-1541) x STRP_STRQ (1-440)
 QY 33 ATAGCTGCTCGAATGGCTACTAGATCGTCTCTGTAATAAACAAGCCAAATTTGGTTGT 92
 Db 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
 QY 93 ACCGTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTTAAATTTTGAATC 152
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
 QY 153 GATCTAACATCAGCACCTGCTCATGGAGGAAGAGACAGCAAGCGCTTAAGTCAAAATCA 212
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 213 AAACCATTTGCTAGTAGTGGCCGATGTCACATAACTTGAGAAAGCTGACTACTA 272
 Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
 QY 273 AAGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC 332
 Db 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126

QY 333 ATTGATTTTGAAGCGATGCAACATTACTGATCGAAACGCGAGGTCTACTTTGCTGAC 392
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 393 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 452
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnGluPheLeuLeuSerGlyHis 166
 QY 453 GTGGCGTTAGACCATATAAAGAAACCAATACAAACCAACCAAGCAATCTGTTGATGTG 512
 Db 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 572
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 632
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 633 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 692
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 693 TCCTCAATCGTCACTCATGACATGATTTTCGTCAGATTTTACCAATGGATCAAGAG 752
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 753 TTTTACTTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
 Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 813 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAACGTCTTAAAAAAGGG 872
 Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAATGTTTCCACCAATACGTT 932
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 933 GATGTCGATACCAACGAATTTGCTAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 992
 Db 327 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGTAAGCTAAAGCTACTCTACAAATCTC 1052
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 1112
 Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1113 ACCAACCGTATACCAACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGTAC 1172
 Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1173 CATTTAGCT 1181
 Db 407 HisLeuAla 409
 RESULT 2
 Q53284 STREQ PRELIMINARY; PRT; 414 AA.
 ID Q53284;
 AC Q53284;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE SKC-2.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
RA Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,
RA Rubiera R., Pedraza A., Padron G., Antuch W., la Fuente J., et al.;
RT "High level expression of streptokinase in *Escherichia coli*.";
RL Biotechnology 0:1138-1142 (1992).
DR EMBL; S46536; AAC60418.1; -; Genomic_DNA.
DR PDB; 1C4P; X-Ray; A/B/C/D=149-285.
DR SMR; Q53284; 12-372.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
SQ SEQUENCE 414 AA; 47254 MW; F75BE5831B766904 CRC64;
Alignment Scores:
Pred. No.: 2,61e-131 Length: 414
Score: 1955.00 Matches: 377
Percent Similarity: 98.96% Conservative: 2
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 71.98% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x Q53284_STREQ (1-414)
QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTCTGTAAATAACAGCAATGGTGT 92
DB 1 MetAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 93 AGCGTTGCTGTTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluLe 40
QY 153 GATCTAACTACAGACCTCTCATGAGGAGGAAGACAGACAGCGTCTAAGTCCAAATCA 212
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATATACTTGAGAAAGCTGACTACTA 272
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 273 AAGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAAGACGACTACTTTGAGGTC 332
DB 81 LysAlaIleGlnGlnLeuLeuAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 333 ATTGATTTTCAAGCGATGCAACCAATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 392
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 393 AAAGATGGTTTCGGTAACCTTGGCGACCCCACTGCTCCAGAAATTTTGTAAAGCGACAT 452
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 453 GTGGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATGG 512
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 513 GAATATAGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 573 GATCTAAGCTATTGAAACACTAGCTATCGGTGACCATCATCATCTCAAGATTTACTA 632
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTTTATGACGTGAC 692
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 693 TCCTCAATCGTCACATCATGACATGACATTTTCGCTACGATTTTACCAATGGATCAAGAG 752
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
753 TTACTTTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 812
DB 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
QY 813 AATGAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 872
DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 873 GAAAGCCGCTATGATCCCTTTGATCGCAGTCACTTGAACCTTCAACCTCCACCAATACGTT 932
DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 933 GATGTCGATACCAACGAATTCCTAAAAAGTACGAGCTCTTAAACAGCTAGCGAACGTAAC 992
DB 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGATGAAGCTAAAGCTAACTACTCTACAACATCTC 1052
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1053 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1112
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1113 ACCAACCGTATACACCGTTTATATCGCAAGCGACCCGAGGAGAGAGATGCTAGCTAC 1172
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1173 CATTTAGCT 1181
DB 381 HisLeuAla 383
RESULT 3
Q6UK57_STREQ
ID Q6UK57_STREQ PRELIMINARY; PRT; 436 AA.
AC Q6UK57;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Streptokinase.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35666;
RA Costa C.S., Torres F.A.G., Filho S.A.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY368335; AAQ73571.1; -; Genomic_DNA.
DR SMR; Q6UK57; 59-394.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 436 AA; 49658 MW; CA2DD95F957D8F3C CRC64;
Alignment Scores:
Pred. No.: 1.9e-130 Length: 436
Score: 1943.00 Matches: 374
Percent Similarity: 98.43% Conservative: 3
Best Local Similarity: 97.65% Mismatches: 6
Query Match: 71.54% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x Q6UK57_STREQ (1-436)
QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGTCTCTGTAAATAACAGCAATGGTGT 92

Db 23 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 42
 Qy 93 AGCGTTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
 Db 43 SerValAlaGlyThrValGluGlyThrAsnGlnAspPheLeuLeuPheGluIle 62
 Qy 153 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGACAAAGCTTAAAGTCCAAATCA 212
 Db 63 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 82
 Qy 213 AAACATTGCTGATGATGAGGCGGATGTACATAAATCTGAGAAAGCTGACTACTA 272
 Db 83 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 102
 Qy 273 AAGGCTATTCAAGACCAATTGATCCCTAACGTCACAGTAAAGCAGCACTACTTGAAGTC 332
 Db 103 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 122
 Qy 333 ATTGATTTCGAAGCGATGCAACCAATTACTGATCGAAACGCAAGTCTACTTGTGAC 392
 Db 123 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIysValTyrPheAlaAsp 142
 Qy 393 AAAGATGTTGCGTAACCTTCGCGACCCCACTGTCCAAAGAATTTTGTAAAGCGGACAT 452
 Db 143 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 162
 Qy 453 GTGCGCTGACACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATG 512
 Db 163 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 182
 Qy 513 GAATATCTGTCAGTCTTACTCCCTTAAACCCCTGATGACGATTCAGACGAGTCTCAA 572
 Db 183 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 202
 Qy 573 GATACTAAGCTATTGAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTACTA 632
 Db 203 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 222
 Qy 633 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAGTGAC 692
 Db 223 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 242
 Qy 693 TCCTCAATCGTCACCTCATGACAATGACATTTTCGTGACGATTTTACCAATGATCAAGAG 752
 Db 243 SerSerIleValThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGlu 262
 Qy 753 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTG 812
 Db 263 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 282
 Qy 813 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAGG 872
 Db 283 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 302
 Qy 873 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTTTTCCACCATCAAAATCGTT 932
 Db 303 GluLysProTyrAspProPheAspArgSerHisLysLysLeuPheThrIleLysTyrVal 322
 Qy 933 GATCTGATACCAACGAATTGCTAAAAAGTGAAGAGCTCTTAAACAGCTAGCCAAAGTAA 992
 Db 323 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 342
 Qy 993 TTAGACTTCAGAGATTATACATCTCTGATGATAAGGCTTAAACTACTCTACCAACATCTC 1052
 Db 343 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 362
 Qy 1053 GATGCTTTTGGTATTATGACTATACCTTAACCTGAAAAAGTGAAGATAATCACCATGAC 1112
 Db 363 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 382
 Qy 1113 ACCAACCGTATCATACCGTTTATATGGCGAAGCGACCGACCGAGAGAGATCTAGCTAC 1172

Db 383 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 402
 Qy 1173 CATTAGCT 1181
 Db 403 HisLeuAla 405
 RESULT 4
 STRP_STRS1
 ID STRP_STRS1 STANDARD; PRT; 440 AA.
 AC P10519;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Streptokinase G precursor.
 GN Name=skg;
 OS Streptococcus sp. (strain 19909).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=69017;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RX MEDLINE=89160285; PubMed=2922269;
 RA Walter F., Siegel M., Malke H.;
 RT "Nucleotide sequence of the streptokinase gene from a group-G
 Streptococcus";
 RL Nucleic Acids Res. 17:1262-1262(1989).
 CC -!- FUNCTION: This protein is not a protease, but it activates
 plasminogen by complexing with it. As a potential virulence
 factor, it is thought to prevent the formation of effective fibrin
 barriers around the site of infection, thereby contributing to the
 invasiveness of the cells.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 CC EMBL; X13400; CAA31766.1; -; Genomic_DNA.
 DR PIR; S02723; S02723.
 DR HSP; P00779; 1QOR.
 DR SMR; P10519; 63-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Plasminogen activation; Signal; Virulence.
 FT SIGNAL 1 26
 FT CHAIN 27 440 Streptokinase G.
 SQ SEQUENCE 440 AA; 50199 MW; 5521P8825FE1B6EA CRC64;

Alignment Scores:
 Pred. No.: 2,24e-130 Length: 440
 Score: 1942.00 Matches: 375
 Percent Similarity: 98.43% Conservative: 2
 Best Local Similarity: 97.91% Mismatches: 6
 Query Match: 71.50% Indels: 0
 DB: Gaps: 1

US-09-940-235-9 (1-1541) x STRP_STRS1 (1-440)

Qy 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATGGTTGTT 92
 Db 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
 Qy 93 AGCGTTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspLysSerLeuLysPhePheGluIle 66
 Qy 153 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 212
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86

QY	213	AAACCAATTTGCTACTGATAGTGGCGGATGTGCACATAAACTTGAGAAAGCTGACTTACTA	272	
Db	87	LeuLeuPheAlaThrAspSerGlyAlaMetProHisLeuLeuGluLysAlaAspLeuLeu	106	
QY	273	AAAGGCTATTCAAGAACAAATTTGATCGCTACAGTCCACAGTAAACGACGACTACTTTGAGGTC	332	
Db	107	LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	126	
QY	333	ATTGATTTTGCAGCGATCGAACCAATTACTGATCGAAACGGCAAGGTCTACTTTTGTCTGAC	392	
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	146	
QY	393	AAAGATGGTTTCGGTAACTTGGCGGACCCAACTGTCCAAAGATTTTGTCTAAGCGGACAT	452	
Db	147	LysAspGlySerValThrLeuProIleGlnProValGlnGluPheLeuLeuLysGlyHis	166	
QY	453	GTGCGCGTTAGACCATATAAGAAAGAAACCAATACAAACCAACGCGGAATCTGTTGATGTG	512	
Db	167	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186	
QY	513	GAATATACTGTACAGTTTACTCCCTTAAACCTCGATGACGATTTGACACCGGTCTCTCAA	572	
Db	187	GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProAlaLeuLys	206	
QY	573	GATCTAAGACTATTGAAACACACTAGCTATCGGTGACCCATCACATCTCAAGAAATTACTA	632	
Db	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu	226	
QY	633	GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTATGAAACGTGAC	692	
Db	227	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	246	
QY	693	TCCTCAATCGTCACATCATGACAAATGACATTTTCCGTACGATTTTCAATGATCAAGAG	752	
Db	247	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	266	
QY	753	TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTCGTCTG	812	
Db	267	PheThrTyrHisValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	286	
QY	813	AATGAGAATAATAACAACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG	872	
Db	287	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	306	
QY	873	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTCCACCATCAAAATACGTT	932	
Db	307	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326	
QY	933	GATGTTCGATACCAACGAATTTGCTAAAAAGTAGGACGAGCTCTTAAACAGTAGCGAAACGTAAAC	992	
Db	327	AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346	
QY	993	TTAGACTTCAGAGATTTATACATCCTCGTGATAGGCTAAACTACTCTACAAACAATCTC	1052	
Db	347	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	366	
QY	1053	GATGCTTTTGGTATTATATGCACTATACCTTAACTGGAAAAGTAGAGGATAATCACCATGAC	1112	
Db	367	AspAlaPheGlyIleMetAspPyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	386	
QY	1113	ACCAACCGTATCATACACGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAC	1172	
Db	387	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	406	
QY	1173	CATTAGCT 1181		
Db	407	HisLeuAla 409		
RESULT 5				
Q7X0Y1_STRPY				
ID	Q7X0Y1_STRPY	PRELIMINARY;	PRT;	440 AA.
AC	Q7X0Y1;			
DT	01-OCT-2003 (TrEMBLRel. 25, Created)			

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D489;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234137; AAP39957.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y1; 63-398.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0008243; P:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; S:Staphylokinase.
DR InterPro; IPR008124; S:Streptokinase.
DR Pfam; PF02821; S:Staphylokinase; 3.
DR PRINTS; PR01753; S:STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49903 NW; CC4E6E9647043BAC CRC64;

Alignment Scores:
Pred. No.: 3.25e-121 Length: 440
Score: 1814.00 Matches: 347
Percent Similarity: 95.04% Conservative: 17
Best Local Similarity: 90.60% Mismatches: 19
Query Match: 66.79% Indels: 0
DB: 2 Gaps: 0
US-09-940-2335-9 (1-1541) x Q7X0Y1 STRPY (1-440)

```

Qy	33	ATAGCTGGTCTCGAATGCTACTAGATCGTCTCTTCTGTAAATAACAGCAATTTGGTTGTT	92
Db	27	IleAlaGlyIyrGlyTrpLeuProAspArgProProlIeAsnAsnSerGlnLeuValVal	46
Qy	93	AGCGTTGCTGTTACTGTTTGAGGGAGCAATCAAGACATTAGTCTTAAATTTTTCGAATC	152
Db	47	SerMetAlaGlyIleValGlnGlyThrAspLysLysValPheIleAsnPheGluIle	66
Qy	153	GATCTAAATCATCAGCACTGCTCATCGAGGAAGACAGACGAGGCTTAGTCCAAATATCA	212
Db	67	AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	86
Qy	213	AAACCATTTGCTACTGATAGTGGCGCGAGTGCACATAAATTTGAGAAAGCTGACTTACTA	272
Db	87	LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106
Qy	273	AAGGCTATTCAAGAACAAATTGATCGCTACGCTCCACAGTCAACGACGACTACTTTGAGTC	332
Db	107	LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal	126
Qy	333	ATTGATTTTGCACGATGCAACCACTTACTGATCGAAACGGCAAGCTTACTTTGCTGAC	392
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyGlnValTyrPheAlaAsp	146
Qy	393	AAAGATGTTTCGGTAACTTGGCCGACCAACTGTCTCCAAGAAATTTTTCCTAAGCGGACAT	452
Db	147	LysAspGlySerValThrLeuProThrGlnProValGlnPheLeuLeuSerGlyHis	166
Qy	453	GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGGAATCTGTTGATGTG	512
Db	167	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186
Qy	513	GAATACTACTGTACAGTTTACTTCCCTTAAACCTCGATGACGATTTTCAGACCAAGTCTCAAA	572

RESULT	5
Q7X0Y1	S
ID	Q7X
AC	Q7X
DT	01-


```
Db 187 LysTyrThrValGlnPheThrProLeuAenProAspAspPheArgProGlyLeuLys 206
QY 573 GATACCTAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTTATGAACTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAenLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAenAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATAAATAATCTGCTCG 812
Db 267 PheThrTyrArgValLysAspArgGluGlnAlaTyrGlyLeaenLysLysSerGlyLeu 286
QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACCTCTTAAAAAGGG 872
Db 287 AsnGluGluIleAsnAenThrAspLeuIleSerGluLysTyrTyrIleLeuLysLysGly 306
QY 873 GAAAGCCGTATGATCCCTTTGATCGCACTGCTGAACTGTTCAACATCAAAATACGTT 932
Db 307 GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCGATACCAACGAATTGCTTAAAGTGGAGCTCTTAACAGCTAGCGCAACGTAAC 992
Db 327 AspValAenThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAen 346
QY 993 TTAGACTTTCAGAGATTATACGATCTCTGATGATGAGCTTAACTACTCTCAACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAenLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTACAGGATATACAGATGAC 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACGTCATCAATACCGTTTATATGGCAAGCGACCGAGGAGAGATGCTAGCTAC 1172
Db 387 AsnAenArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409

RESULT 6
QY0Y8 STRPY PRELIMINARY; PRT; 440 AA.
AC QY0Y8;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=88-019;
RA PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RX Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234129; AAP39949.1; -; Genomic_DNA.
DR HSPF; Q53284; IC4P.
DR SMR; Q7X0Y8; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
```

```
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREP_KINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49885 MW; 69DB44F4026E3975 CRC64;

Alignment Scores:
Pred. No.: 2 34e-120 Length: 440
Score: 1802.00 Matches: 347
Percent Similarity: 94.52% Conservative: 15
Best Local Similarity: 90.60% Mismatches: 21
Query Match: 66.35% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q7X0Y8_STRPY (1-440)
QY 33 ATAGCTGGTCTCGAATGGCTACTAGATCGCTCTTCTGTAAATAACACCCAAATGGTTGTT 92
Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProValAenAenSerGlnLeuVal 46
QY 93 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTAATCATCACGACCTGCTCATGAGGAAAGACAGCAAGCTTAAAGTCCAAATCA 212
Db 67 AspLeuThrSerGlnHisAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAACTTGAGAAAGCTGACTACTA 272
Db 87 LysProPheAlaThrAspAenGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AGGCTATTCAAGAACAAATGATCGCTTAAGCTCCACAGTAACGACGACTACTTTGAGGTC 332
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAenValHisSerAenAspGlyTyrPheGluVal 126
QY 333 ATTGATTTCGACAGCATGCAACCATTAATGTCGAAACCGCAAGGTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAGATGTTTCGGTAACTTCCTTAAACCCCTGATGACGATTTTCAGCCAGGCTCTCAA 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 453 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCCAGGAAATCTGTTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAenGlnAlaLysSerValAspVal 186
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCCAGGCTCTCAA 572
Db 187 LysTyrThrValGlnPheThrProLeuAenProAspAspPheArgProGlyLeuLys 206
QY 573 GATACCTAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTTTATGAACTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAenLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAenAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATAAATAATCTGCTCG 812
Db 267 PheThrTyrArgValLysAspArgGluGlnAlaTyrGlyLeaenLysLysSerGlyLeu 286
QY 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 872
Db 287 AsnGluGluIleAsnAenThrAspLeuIleSerGluLysTyrTyrIleLeuLysLysGly 306
```

Qy	873	GAAAGCGGTATGATCCCTTTGATCGAGTCACCTCTGAAACACTGTTACCATCAAAATACGTT	933
Db	307	GlusSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326
Qy	933	GATGTGCAGTACCAACGAATTCTCTAAAGTCAGCAGCTCTTAAACAGCTAGCGAACGTAAC	992
Db	327	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346
Qy	993	TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAACAATCTC	1052
Db	347	LeuAspPheArgAspLeuTyrAspProCysAspLysAlaLysLeuLeuTyrAsnAsnLeu	366
Qy	1053	CATGCTTTTGGTATTATGCGACTATACCTTAAGTAAAGTAGAGGATAATACAGATGAC	1112
Db	367	AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys	386
Qy	1113	ACCAACCGTATCATCAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAC	1172
Db	387	AsnAsnArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr	406
Qy	1173	CATTTAGCT 1181	
Db	407	HisLeuAla 409	
RESULT 7			
Q8KSR8_STRP3	Q8KSR8_STRP3 PRELIMINARY; PRT; 440 AA.		
DT	Q8KSR8; Q79W73; (TremBLrel. 22, Created)		
DT	01-OCT-2002 (TremBLrel. 22, Last sequence update)		
DT	01-OCT-2002 (TremBLrel. 22, Last sequence update)		
DT	13-SEP-2005 (TremBLrel. 31, Last annotation update)		
DE	Streptokinase A.		
GN	Name=ska; OrderedLocusNames=Spel1700, SpyM3_1698;		
OS	Streptococcus pyogenes (serotype M3).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=301448;		
RN	(1)		
RN	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=MGAS315 / Serotype M3;		
RC	MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;		
RA	Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,		
RA	Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,		
RA	Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,		
RA	Schlievert P.M., Mueser J.M.;		
RT	"Genome sequence of a serotype M3 strain of group A Streptococcus:		
RT	phase-encoded toxins, the high-virulence phenotype, and clone		
RT	emergence.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).		
RN	(2)		
RN	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=SSI-1 / Serotype M3;		
RC	MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;		
RA	Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,		
RA	Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,		
RA	Hayashi H., Hattori M., Hamada S.;		
RT	"Genome sequence of an M3 strain of Streptococcus pyogenes reveals a		
RT	large-scale genomic rearrangement in invasive strains and new insights		
RT	into phase evolution.";		
RL	Genome Res. 13:1042-1055(2003).		
RL	EMBL; AA014169; AAM80305.1; -; Genomic_DNA.		
DR	EMBL; BA000034; BAC64795.1; -; Genomic_DNA.		
DR	HSP; P00779; 1L4D.		
DR	SMR; Q8KSR8; 63-398.		
DR	GO; GO:0016301; F:Kinase activity; IEA.		
DR	GO; GO:0008243; F:Plasminogen activator activity; IEA.		
DR	GO; GO:0005515; F:Protein binding; IEA.		
DR	InterPro; IPR004093; Staphylokinase.		
DR	InterPro; IPR008124; Streptokinase.		
DR	Pfam; PF02821; Staphylokinase; 3.		
DR	PRINTS; PR01753; STREPKINASE.		
DR	Complete proteome; Kinase.		
SW	SEQUENCE 440 AA; 49911 MW; 7CC644F4026E3975 CRC64;		

Alignment Scores:			
Pred. No.:	2,34e-120	Length:	440
Score:	1802.00	Matches:	347
Percent Similarity:	94.52%	Conservative:	15
Best Local Similarity:	90.60%	Mismatches:	21
Query Match:	66.35%	Indels:	0
DB:	2	Gaps:	0
US-09-940-235-9 (1-1541) x Q8KSR8_STRP3 (1-440)			
QY	33	ATAGCTGGTCCCTGAATGGCTACTAGATCGCTCTCTGTAAATAACAGCCAAATTTGGTGT	92
DB	27	IleAlaGlyThrGlyTrpLeuProAspArgProValAsnAsnSerGlnLeuVal	46
QY	93	AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC	152
DB	47	SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle	66
QY	153	GATCTAACATCAGCAGCTGCTCATGGAGAAAGACGAGCARGGCTTAGTCCAAATCA	212
DB	67	AspLeuThrSerGlnHisAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer	86
QY	213	AAACCATTTGCTACTGATAGTCGCGGAGTGCACATAAACTTGAGAAAGCTGACTTACTA	272
DB	87	LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106
QY	273	AAGGCTATTCAAGAACAATTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC	332
DB	107	LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal	126
QY	333	ATTGATTTTGCNAGCGATGCAACCATTAAGTATCGAATCGAAACGGCAGGTCTACTTTGCTGAC	392
DB	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	146
QY	393	AAAGATGGTTCCGTAACCTTGCCGACCCCAACCTGTCCAAGATTTTGTAAAGCGGACAT	452
DB	147	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	166
QY	453	GTGCGCGTTAGACCATATAAGAAAAACAATACAAACCAAGCGAAATCTGTTGATGTG	512
DB	167	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186
QY	513	GAATATCTGTACAGTTACTCCCTTAACCTGATGACGATTCAGACCAAGTCTCAA	572
DB	187	LysTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys	206
QY	573	GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA	632
DB	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	226
QY	633	GCTCAAGCACAAAGCAATTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC	692
DB	227	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	246
QY	693	TCCTCAATCGTCACATGACAAATGACATTTCCGTACGATTTTACCATTGATCAAGAG	752
DB	247	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	266
QY	753	TTTACTTACCGTGTAAAAATTCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTCG	812
DB	267	PheThrTyrArgValLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyLeu	286
QY	813	AATGAAGAATAAACACACTGACTGATCTCTGAGAAATATATAGCTCTTAAAAAGGG	872
DB	287	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrIleLeuLysLysGly	306
QY	873	GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTCACCATCAATACGTT	932
DB	307	GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326
QY	933	GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGTAGCCGACCTAAC	992

QY 873 GAAAGCCGTATGATCCCTTTGATCGAGTCACCTTGAAACTGTTCCACCATCAATACGTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCATACCAACCAATTTGCTAAAGAGGAGCAGCTCTTAACAGCTAGCGAAGTAAC 992
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCACGATGAC 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACCGATCATACACCTTTATATGGCAAGCGCCGAGAGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
RESULT 9
Q53ZX6_STRPY
ID Q53ZX6_STRPY PRELIMINARY; PRT; 440 AA.
AC Q53ZX6.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86-779.
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234128; AAP39948.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;
Alignment Scores:
Pred. No.: 8.76e-120 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.9% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 66.05% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x Q53ZX6_STRPY (1-440)
QY 33 ATAGCTGGTCTGAATGGCTACTAGATCGTCTCTGTAATTAACAGCCAAATGGTTGTT 92
Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProProIleAsnAsnSerGlnLeuValVal 46
QY 93 AGCGTTGCTGGTACTGTTGAGGGGACCAATCAAGACATTAGTCTTAATATTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTAACATCATGACCTCTCATGAGGAAGACAGACAGCGCTTAAGTCCAAATCA 212
Db 67 AspLeuThrSerGlnProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCAATTGCTACTGATAGTGGCGGATGCTCATATAAATTCAGAAAGCTGACTTACTA 272

Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AAGGCTATTCAAGAACAAATTCATCGCTAACCTCCACAGTAACAGACACTACTTTGAGTTC 332
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTTGAACCGATGCAACCATTAATCTGATCGAAACGGCAGGCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAAAGATGGTTCCGTAACCTTCGCCGACCACTGCTCAAGAAATTTTCTAAGCGGACAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLysGlyHis 166
QY 453 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAAACCAAGCGAATCTGTTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCCAGACGAGTCTCAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATCTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 633 GCTCAAGCAACAACCATTTTAAACAAAACCCAGCGCTATACGATTTATGAACGTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCATCGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCTGGTGTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY 813 AATGAAGAAATAACAACACACTGACTGATCTCTGAGAAATATTAGTCCTTAAAAAAGG 872
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
QY 873 GAAAGCCGTATGATCCCTTTGATCGAGTCACCTTGAAACTGTTCCACCATCAATACGTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 933 GATGTCATACCAACCAATTTGCTAAAGAGGAGCAGCTCTTAACAGCTAGCGAAGTAAC 992
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTGCTGATTATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCACGATGAC 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACCGATCATACACCTTTATATGGCAAGCGCCGAGAGAGAGATGCTAGCTAC 1172
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
RESULT 10
Q7X0Y2_STRPY
ID Q7X0Y2_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y2.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D3067;
RX PubMed=1479231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121 (2004).
DR EMBL: AY234136; AAP39956.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y2; 63-398.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0008243; P:plasmaingogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50188 MW; 4CB42AC2A9062C2E CRC64;
Alignment Scores:
Pred. No.: 2.78e-119 Length: 440
Score: 1787.00 Matches: 342
Percent Similarity: 94.26% Conservative: 19
Best Local Similarity: 89.30% Mismatches: 22
Query Match: 65.80% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x Q7X0Y2_STRPY (1-440)
QY 33 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTCTTCTGTAATAATACAGCAATGGTGTGTT 92
Db 27 IleaAGlyTyrGlyTrrLeuProAspArgProProlIeAsnAsnSerGlnLeuValVal 46
QY 93 AGCGTGTGCTGCTACTGTTTGAGGGAGCAATCAAGACATCTTAGTCTTAAATTTTTCGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTAATCATCAGCACCTGCTCATGAGGAGAAAGACAGACAGCAAGGCTTAAGTCCAAAATCA 212
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCATTTGCTACTGATAGTCGGCGGATGTCACATAACTTGAGAAAGCTGACTTACTA 272
Db 87 LysProPheAlaThrAspAsnSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AAGGCTATTCAAGAACATTCATCGCTAAAGTCCACAGTAAAGCAGCTACTTTTGAGGTC 332
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTTGAACGATGCAACCATTAATCTGATCGAAACGGCAAGGTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 393 AAGATGGTTCGGTAAACCTTCGGACCCAACTGTCTCAAGATTTTTCGTACGCGACAT 452
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnLysPheLeuLeuThrGlyHis 166
QY 453 GTGCGCGTTTACACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTGTATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 513 GAATATATCTGATAGTTTACTCCCTTAAACCTGTAGACGATTTCAGACCGGTCTCAAA 572
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys 206

QY	573	GATCTAAGCTATTGAAACAC	TAGCTATACGGTGACCACT	CACTCATCTCAAGAAATTACTA	633							
Db	207	AspThrIysLeuLeuLysLeu	AlaIleGlyAspThrValThr	SerGlnGluLeuLeu	226							
QY	633	GCTCAAGCACAAAGCATTTT	AAACAAACACCAAGCTATAC	GATTATGACAGCTGAC	692							
Db	227	AlaGlnAlaGlnSerIleLeu	AsnGlnSerHisProAspTyr	ThrIleTyrGluArgAsp	246							
QY	693	TCCTCAATCGTCATCATGA	CAATGATTTTCCGTACGATT	TACCAATGGATCAAGAG	752							
Db	247	SerSerIleValThrHisAsp	LysAspIlePheArgThrIle	LeuProMetAspGlnGlu	266							
QY	753	TTTACTTACCGTGTAAATAA	TCGGAAACAGCTTATAGGAT	CAATAAAAAATCTGGCTG	812							
Db	267	PheThrTyrArgValLysAsn	ArgGluGlnAlaTyrArgIle	AsnLysLysSerGlyLeu	286							
QY	813	AATGAAGAAATAAACACACT	GCCTGATCTCTGAGAAATAT	TACGTCCCTTAAAAAAGGG	872							
Db	287	AsnGluGluIleAsnAsnThr	AspLeuIleSerGluLysTyr	TyrValLeuLysLysArg	306							
QY	873	GAAAGCGCGTATGATCCCTT	TGATCGCAGTCACCTTGAAC	TGTTCCACCATCAAAATACGTT	932							
Db	307	GluLysProTyrAspProPhe	AspArgSerHisLeuLysLeu	PheThrIleLysTyrVal	326							
QY	933	GATGTCGATACCAACGAATT	CTCTAAAAAGTGAGCAGCTC	TTAAACAGCTAGCGAAACGTAAAC	992							
Db	327	AspValAsnThrAsnGluLeu	LysLysSerGluGlnLeuLeu	ThrAlaSerGluArgAsn	346							
QY	993	TTAGACTTCAGAGATTTAT	ACGATCCCTCGTGATAAAGG	TAAAGCTACTCTACAAACATCTC	1052							
Db	347	LeuAspPheArgAspLeuTyr	AspProArgAspLysAlaLys	LeuLeuTyrAsnAsnLeu	366							
QY	1053	GATGCTTTTGGTATTATGG	ACTATACCTTAACTCGAAAA	AGTAGAGGATAATCACGATGAC	1112							
Db	367	AspAlaPheAspIleMetAsp	TyrThrLeuThrGlyLysVal	GluLysAspAsnHisAspLys	386							
QY	1113	ACCAACCGTATCATAAACG	TTTATATGGCAAGCGACCC	GAAAGGAGAGAATGCTAGCTAC	1172							
Db	387	AsnAsnArgValValThrVal	TyrMetGlyLysArgProLys	GlyAlaLysGlySerTyr	406							
QY	1173	CATTTAGCTT1181										
Db	407	HisLeuAla409										
RESULT 11												
ID	Q7X0Y7_STRPY	PRELIMINARY; PRT; 440 AA.										
AC	DT	Q1-OCT-2003 (TrEMBLrel. 25, Created)										
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)											
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)											
DE	Streptokinase.											
GN	Name=ska;											
OS	Streptococcus pyogenes.											
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;											
OC	Streptococcus.											
OX	NCBI_taxID=1314;											
RN	[1]											
RP	NUCLEOTIDE SEQUENCE.											
RC	Strain=1RP112;											
RX	PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;											
RA	Kalia A., Bessen D.E.;											
RT	"Natural selection and evolution of streptococcal virulence genes involved in tissue-specific adaptations.";											
RL	J. Bacteriol. 186:110-121(2004).											
DR	EMBL; AY234130; AAP39950.1; -; Genomic_DNA.											
DR	HSP; Q53284; 1C4P.											
DR	SMR; Q7X0Y7; 63-398.											
DR	GO; GO:0016301; P:kinase activity; IEA.											
DR	GO; GO:0008243; F:plasminogen activator activity; IEA.											
DR	GO; GO:0005515; F:protein binding; IEA.											
DR	InterPro; IPR004093; Staphylokinase.											
DR	InterPro; IPR008124; Streptokinase.											

DR Pfam: PF02821; Staphylokinase; 3.
DR PRINTS: PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 3.3e-117 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservative: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 64.73% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q7X0Y7_STRPY (1-440)

QY 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGCTCTCTGTAATAACAGCCAAATGGTTGTT 92
DB 27 IleAlaGlyTyrGlyTyrLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 93 AGCGTTGCTGACTCTTGGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTTTGAATC 152
DB 47 SerMetAlaGlyLeuValGluGlyThrAspLysLysValPheIleAsnPheGluLe 66

QY 153 GATCTAATCATGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 212
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86

QY 213 AAACCATTTGCTACTGATGCGCGGATGCATCAATAAAGCTTCAGAAAGCTGACTTACTA 272
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106

QY 273 AAGGCTATTCAAGAACAAATGATCGCTAACGTCACAGTCAACGACGACTACTTTGAGTC 332
DB 107 LysAlaIleGlnGluGluLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

QY 333 ATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 392
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 393 AAAGATGTTTCGGTAACTTCGCGACCACTGTCACAGAAATTTTGTAAAGCGGACAT 452
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166

QY 453 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAACGAAATCTGTTGATGTG 512
DB 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186

QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206

QY 573 GATCTAGCTATTGAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 632
DB 207 AsnThrLysLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226

QY 633 GCTCAAGCAACAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACCTGAC 692
DB 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 752
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 753 TTACTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286

QY 813 AATGAAGAAATAAACHACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 872
DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrValLeuLysLysGly 306

QY 873 GAAAAGCCGTATGATCCCTTTGATCCAGTCACTTCGAACTGTTCAACCATCAATAGCTT 932

DB 93 AGCGTTGCTGACTGTTGAGGGGACCAATCAAGACATTTAGTCTTAAATTTTTTGAATC 152

QY 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326

QY 933 GATGTCGATACCAAGCAATTCCTAAAGAGTCCAGCAGCTCTTAAACAGCTAGCGCAACGTAAC 992

DB 327 AspValAsnThrAsnLysLeuLeuLysSerGluGlnLeuLeuThrIleAlaSerGluArgAsn 346

QY 993 TTAGACTTCAGAGATTTATATACGATCCTCGTGATAGGCTAAAGCTAACTCTACACAATCTC 1052

DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366

QY 1053 GATGCTTTTGGTATTATGGAATATACCTTAATCTGAAAAAGTAGAGGATAATCAGATGAC 1112

DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386

QY 1113 ACCAACCGTATCATACACCGTTTATATGGCAAGCGACCCGAGGAGAGAGAAATGCTAGCTAC 1172

DB 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406

QY 1173 CATTTAGCT 1181

DB 407 HisLeuAla 409

RESULT 12
Q5X9T6_STRP6
ID Q5X9T6_STRP6 PRELIMINARY; PRT; 440 AA.
AC Q5X9T6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Streptokinase (EC 3.4.-.-).
GN OrderedLocustNames=M6_Spy1692;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS10394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.E.,
RA Voyich J.M., DeLeo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
metagenome: complete genome sequence of a macrolide-resistant serotype
M6 strain.";
RL J. Infect. Dis. 190:727-738 (2004).
DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
DR SRR; Q5X9T6; 63-398.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR04093; Staphylokinase.
DR InterPro; IPR08124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS: PR01753; STREPKINASE.
KW Complete proteome; Hydrolase; Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 3.3e-117 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservative: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 64.73% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q5X9T6_STRP6 (1-440)

QY 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGCTCTCTGTAATAACAGCCAAATGGTTGTT 92

DB 27 IleAlaGlyTyrGlyTyrLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 93 AGCGTTGCTGACTCTTGGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTTTGAATC 152

DB 47 SerMetAlaGlyLeuValGluGlyThrAspLysLysValPheIleAsnPheGluLe 66

QY 153 GATCTAATCATGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 212

DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86

QY 213 AAACCATTTGCTACTGATGCGCGGATGCATCAATAAAGCTTCAGAAAGCTGACTTACTA 272

DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106

QY 273 AAGGCTATTCAAGAACAAATGATCGCTAACGTCACAGTCAACGACGACTACTTTGAGTC 332

DB 107 LysAlaIleGlnGluGluLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

QY 333 ATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 392

DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 393 AAAGATGTTTCGGTAACTTCGCGACCACTGTCACAGAAATTTTGTAAAGCGGACAT 452

DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166

QY 453 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAACGAAATCTGTTGATGTG 512

DB 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186

QY 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGTCTCAAA 572

DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206

QY 573 GATCTAGCTATTGAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 632

DB 207 AsnThrLysLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226

QY 633 GCTCAAGCAACAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACCTGAC 692

DB 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 752

DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 753 TTACTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812

DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286

QY 813 AATGAAGAAATAAACHACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 872

DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrValLeuLysLysGly 306

QY 873 GAAAAGCCGTATGATCCCTTTGATCCAGTCACTTCGAACTGTTCAACCATCAATAGCTT 932


```
QY 393 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCCAAGATTTTGTGAAGCGGACAT 452
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 453 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACGAGTCTCAAA 572
Db ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY 573 GATACCTAAGCTATTGAAAAACACTAGCTATCCGGTGACACCATTCACATCTCAAGAAATTA 632
Db AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
Db AlaGlnAlaGlnSerIleLeuAsnGluSerHisSerAspTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
Db SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 753 TTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
Db PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286
QY 813 AATGAAGAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
Db GluGluLysThrAsnAsnThrAspLeuLysSerGluLysTyrTyrValLeuLysLysGly 306
QY 873 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTAACCTTACCATTCAATACGTT 932
Db GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
QY 933 GATGTGATACCAACGAATTTGCTAAAAAGTGACGAGCTTAAACAGCTAGCGACGTAAC 992
Db AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 993 TTAGACTTCAGAGATTTATACGATCTCTGTGTAAGCTAAACTCTCTACAACAATCTC 1052
Db LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1053 GATGCTTTTGGTATTATGACTATATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1112
Db AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAACCGTATCATACCGTATTATATGGCAAGCGACCGAAGAGAGATGCTAGTAC 1172
Db AsnAsnArgValValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
QY 1173 CATTTAGCT 1181
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 407 HisLeuAla 409
RESULT 14
Q7X0Y3_STRPY
ID Q7X0Y3_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y3_
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ALAB49;
```

```
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234134; AAP39954.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y3; 63-398.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0008243; P:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50342 MW; B3C60BAA50DB2C39 CRC64;

Alignment Scores:
Pred. No.: 1,06e-114 Length: 440
Score: 1723.00 Matches: 327
Percent Similarity: 92.43% Conservative: 27
Best Local Similarity: 85.38% Mismatches: 29
Query Match: 63.44% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-9 (1-1541) x Q7X0Y3_STRPY (1-440)
QY 33 ATAGCTGGTCTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAATTGGTTGTT 92
Db 27 IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
QY 93 AGCGTTCTGCTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 153 GATCTAATCACCACCTCTCATGGAGGAAAGACAGACGAGCTTAAGTCCAAAATCA 212
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 213 AAACCATTTGCTACTGATGTGGCGCATGTGCATATAACTTGAGAAAGCTGACTTACTA 272
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 273 AAGCTATTCAAGAACAAATTGATCGCTTAACGTCCACAGTAAACGACGACTACTTTGAGTC 332
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 333 ATTGATTTTGCAGCGATGCAACCATTAAGTATGATCGAAACGCAAGGCTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
QY 393 AAAGATGTTTCGGTAACCTTGGCCGACCAACCTGTCCAAGAAATTTTGTGAAGCGGACAT 452
Db 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
QY 453 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
QY 513 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACGAGTCTCAAA 572
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY 573 GATACCTAAGCTATTGAAAAACACTAGCTATCCGGTGACACCATTCACATCTCAAGAAATTA 632
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 633 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246
QY 693 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 752
```

```
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
Qy 753 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGlnAsnAspAsnLysThrGlyLeu 286
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 872
Db 287 LysLysGluThrLysAsnThrAspLeuIleSerGluLysTyrTyrlleLeuLysLysGly 306
Qy 873 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACCATCAAAATACGTT 932
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Qy 933 GATGTCGATACCAACGAATTCCTTAAAGTGGAGCTCTTAAACAGTACCGACCTAAC 992
Db 327 AspValAspThrLysAspLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Qy 993 TTAGACTTCAGAGATTATACGATCCTCGTGAAGGCTAAACTACTCTACACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Qy 1053 GATGCTTTGGTATTATGACTATATACCTTAACCTGGAAAAAGTAGAGGATAATCACGATGAC 1112
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Qy 1113 ACCAACCGTATCATACCTTTATATGGCAAGCAGCGAGGAGAGATGCTACTAC 1172
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
Qy 1173 CATTTAGCT 1181
Db 407 HisLeuAla 409
```

RESULT 15

```
Q7X0Y0 STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y0*
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=d633; DOI=10.1128/JB.186.1.110-121.2004;
RX PubMed=14679231;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
  involved in tissue-specific adaptations.";
RL J. Bacteriol. 186.110-121(2004).
DR EMBL; AY234138; AAP39958.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y0; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50143 MW; 4ACB2C29349D680C CRC64;
```

```
Alignment Scores:
Pred. No.: 3.95e-114 Length: 440
Score: 1715.00 Matches: 325
Percent Similarity: 91.91% Conservative: 27
```

```
Best Local Similarity: 84.86% Mismatches: 31
Query Match: 63.14% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-9 (1-1541) x Q7X0Y0_STRPY (1-440)
Qy 33 ATAGCTGGTCTGAATGGCTACTAGATCGTCTTCTGTTAAATAACGCCAATTTGGTTGTT 92
Db 27 IleAlaGlyProGluThrLeuLeuGlyArgProProValAsnAsnSerGlnLeuValVal 46
Qy 93 AGCGTTGCTGTTACTGTTGAGGGAGCAATCAAGACATTAAGCTTAAATTTTTTGAATC 152
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
Qy 153 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGCAAGCTTAAGTCCAAATCA 212
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGlnGlnLysLeuSerProLysSer 86
Qy 213 AAACCATTTTCTACTGATGCGCGGATGTCATAAACTTGAAGAAAGCTGACTTACTA 272
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Qy 273 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCCACAGTAAACGACGACTACTTTGAGCT 332
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Qy 333 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACGCGAAGTCTACTTTGCTGAC 392
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
Qy 393 AAAGATGGTTCGGTAACTTCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 452
Db 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnGlnPheLeuLeuArgGlyHis 166
Qy 453 GTCCGCTTAGACCATATAAAGAAAAACCAATACAAACCAACGCGAAATCTGTGATGTG 512
Db 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
Qy 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
Qy 573 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTA 632
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAsnThrIleThrSerGlnGlnLeuLeu 226
Qy 633 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATAGGATCAATAAAAAATCTGGCTG 692
Db 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyrThrIleTyrGluArgAsp 246
Qy 693 TCCTCAATCTGCTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 752
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
Qy 753 TTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGCTG 812
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspIle 286
Qy 813 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 872
Db 287 LysGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
Qy 873 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACCATCAAAATACGTT 932
Db 307 GluGluProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Qy 933 GATGTCGATACCAACGAATTCCTTAAAGTGGAGCTCTTAAACAGTACCGACCTAAC 992
Db 327 AspValAspThrLysAlaLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Qy 993 TTAGACTTCAGAGATTATACGATCCTCGTGAAGGCTAAACTACTCTACACAATCTC 1052
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
```

```

QY 1053 GATGCTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAGGATAATCAGCATCAC 1112
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1113 ACCAAGCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGAGAGAGATGCTAGCTAC 1172
Db ||||||| :||||: ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
387 AsnAsnArgValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1173 CATTAGCT 1181
Db ||||||| |||||||
407 HisLeuAla 409

```

Search completed: January 28, 2006, 02:33:35
 Job time : 183.071 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:53:18 ; Search time 8444.12 Seconds
(without alignments)
11181.386 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gaacacccgcagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_in:
3: gb_env:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pr:
9: gb_ro:
10: gb_ste:
11: gb_sy:
12: gb_un:
13: gb_vi:
14: gb_hlg:
15: gb_pl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150	69.2	2385	6	AR143998 Sequence
2	1148.6	69.2	1242	6	AR144000 Sequence
3	1148.6	69.2	1401	6	I05204 Sequence 5
4	1148.6	69.2	2568	1	STRSKC
5	1148.6	69.2	2568	6	A04926 Streptococ
6	1148.6	69.2	8931	1	SEDEXB
7	1143.8	68.9	2568	6	E00522 S.equisimil
8	1136.2	68.4	7057	6	E00522 DNA fragmen
9	1134.6	68.3	2566	6	CQ97820 Sequence
10	1127.8	67.9	1245	1	AR068768 Sequence
11	1127.6	67.9	1257	1	S46536 SKC-2=strep
12	1127.6	67.9	1257	6	A20015 SEQ ID NO:
13	1126.2	67.8	1407	6	I13203 Sequence 25
14	1126.2	67.8	1512	6	E01413 DNA sequenc
15	1126.2	67.8	1512	6	A20016 SEQ ID NO:
16	1124.6	67.7	1317	6	I13204 Sequence 27
17	1124.6	67.7	1317	6	A20009 SEQ ID NO:
18	1124.6	67.7	1335	6	I13197 Sequence 18
					A20006 SEQ ID NO:

19	1124.6	67.7	1335	6	I13194	I13194 Sequence 14
20	1124.6	67.7	1458	6	A20027	A20027 SEQ ID NO:
21	1124.6	67.7	1458	6	I13215	I13215 Sequence 42
22	1124.6	67.7	1467	6	A20030	A20030 SEQ ID NO:
23	1124.6	67.7	1467	6	I13218	I13218 Sequence 46
24	1124.6	67.7	2588	6	A20021	A20021 SEQ ID NO:
25	1124.6	67.7	2589	6	I13209	I13209 Sequence 34
26	1116.6	67.2	1473	1	SGSKG	X13400 Streptococ
27	1115	67.1	1311	1	AX368335	AX368335 Streptoco
28	1096.6	66.0	1122	6	ARI75893	ARI75893 Sequence
29	1096.6	66.0	1122	6	AX030317	AX030317 Sequence
30	1096.6	66.0	1158	6	ARI75894	ARI75894 Sequence
31	1096.6	66.0	1158	6	AX030318	AX030318 Sequence
32	1089.2	65.6	1209	6	ARI75891	ARI75891 Sequence
33	1089.2	65.6	1209	6	AX030315	AX030315 Sequence
34	1089.2	65.6	1245	6	ARI75892	ARI75892 Sequence
35	1089.2	65.6	1245	6	AX030316	AX030316 Sequence
36	1080.4	65.0	2253	6	I13213	I13213 Sequence 39
37	1078.8	64.9	1119	6	I13206	I13206 Sequence 30
38	1078.6	64.9	2252	6	A20025	A20025 SEQ ID NO:
39	1066.8	64.2	1118	6	A20018	A20018 SEQ ID NO:
40	1012.6	61.0	1323	1	AY234129	AY234129 Streptoco
41	1012.6	61.0	50354	1	AE014169	AE014169 Streptoco
42	1012.6	61.0	110000	1	BA000034_16	Continuation (17 o
43	1012.6	61.0	110000	1	BA000034_17	Continuation (18 o
44	1011	60.9	1323	1	AY234137	AY234137 Streptoco
45	1004.6	60.5	1323	1	AY234136	AY234136 Streptoco

ALIGNMENTS

RESULT 1
LOCUS ARI43998 2385 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6210667.
ACCESSION ARI43998
VERSION ARI43998.1 GI:15105865
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2385)
AUTHORS Reed,G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 1 03-APR-2001;
FEATURES Location/Qualifiers
source 1. 2385
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	69.2%;	Score 1150;	DB 6;	Length 2385;
Best Local Similarity	99.6%;	Pred. No. 8.3e-292;		
Matches 1153;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
Qy	180	ATGATTCTGGACTGAGTGGTCTAGACCGTCCATCTGTCAACACAGCAATTGGTT	239	
Db	1141	AGGATTCTGGACTGAGTGGTCTAGACCGTCCATCTGTCAACACAGCAATTGGTT	1200	
Qy	240	GTTAGCGTTCTGTTACTGTTGAGGGACGAGTCAAGACATAGTCTTAATTTTGGAA	299	
Db	1201	GTTAGCGTTCTGTTACTGTTGAGGGACGAGTCAAGACATAGTCTTAATTTTGGAA	1260	
Qy	300	ATCGATCTAACATCAGCACTGCTCTGAGGAGGAAGACAGCAAGCTTAAGTCCAAAA	359	
Db	1261	ATCGATCTAACATCAGCACTGCTCTGAGGAGGAAGACAGCAAGCTTAAGTCCAAAA	1320	
Qy	360	TCAAAACCAATTGCTACTGATAGTGGCGCGATGTACATAACTTTGAGAAGCTGACTTA	419	
Db	1321	TCAAAACCAATTGCTACTGATAGTGGCGCGATGTACATAACTTTGAGAAGCTGACTTA	1380	
Qy	420	CTAAGAGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAG	479	

1381 CTAAAGGCTATTCAAGAACTTTGATCGCTAACTCCACAGTAAACGACGACTACTTTGAG 1440
480 GTCAATGATTTTGAAGCGATGCAACCACTACTGATCGAAACGCGAAGTCTACTTTGCT 539
1441 GTCAATGATTTTGAAGCGATGCAACCACTACTGATCGAAACGCGAAGTCTACTTTGCT 1500
540 GACAAAGATGGTTCCGTACCTTGGCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGA 599
1501 GACAAAGATGGTTCCGTACCTTGGCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGA 1560
600 CATGTGCGGTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 659
1561 CATGTGCGGTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 1620
660 GTGGAATATATCTGATCAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTC 719
1621 GTGGAATATATCTGATCAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTC 1680
720 AAGGATATCTAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 779
1681 AAGGATATCTAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1740
780 CTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 839
1741 CTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGT 1800
840 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 899
1801 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 1860
900 GAGTTTACTTACCGTGTAAACATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 959
1861 GAGTTTACTTACCGTGTAAACATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGT 1920
960 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAA 1019
1921 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAA 1980
1020 GGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAATATC 1079
1981 GGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAATATC 2040
1080 GTTGATGTGATACCAACGAATTTGATAAGGTGAGCAGCTCTTAAACAGCTAGCGAACGT 1139
2041 GTTGATGTGATACCAACGAATTTGATAAGGTGAGCAGCTCTTAAACAGCTAGCGAACGT 2100
1140 AACTTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTTAAACTCTTCAACAAT 1199
2101 AACTTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTTAAACTCTTCAACAAT 2160
1200 CTGATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTATAGGATATACGAT 1259
2161 CTGATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTATAGGATATACGAT 2220
1260 GACACCAACCGTATACCAACCGTTTATGCGCAAGCGACCCGAAAGAGAGAAATGCTAGC 1319
2221 GACACCAACCGTATACCAACCGTTTATGCGCAAGCGACCCGAAAGAGAGAAATGCTAGC 2280
1320 TATCAATTTAGCGGTGGT 1337
2281 TATCAATTTAGCGGTGGT 2298

RESULT 2
AR144000
LOCUS AR144000 1242 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6210667.
ACCESSION AR144000
VERSION AR144000.1 GI:15105867
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Reed, G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 5 03-APR-2001;
FEATURES Location/Qualifiers
source 1. .1242
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 69.2%; Score 1148.6; DB 6; Length 1242;
Best Local Similarity 99.7%; Pred. No. 1.9e-291;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTGGTTGTT 242
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTAGTTGTT 60
QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
DB 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 362
DB 121 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 422
DB 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 240
QY 423 AAGGCTATTCAAGAACTGATGCGTAACTGATCGTAACTGATCGTAACTGATCGTAACTGATCGT 482
DB 241 AAGGCTATTCAAGAACTGATGCGTAACTGATCGTAACTGATCGTAACTGATCGTAACTGATCGT 300
QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGTAACTGATCGTAACTGATCGTAACTGATCGT 542
DB 301 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGTAACTGATCGTAACTGATCGTAACTGATCGT 360
QY 543 AAAGATGGTTGCGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 602
DB 361 AAAGATGGTTGCGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 420
QY 603 GTGGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
DB 421 GTGGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY 663 GAATATATCTGATGAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGACAGGCTCTCAA 722
DB 481 GAATATATCTGATGAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGACAGGCTCTCAA 540
QY 723 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 782
DB 541 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 600
QY 783 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAAGCTGAC 842
DB 601 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAAGCTGAC 660
QY 843 TCCTCAATCTGATCATGACAAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 902
DB 661 TCCTCAATCTGATCATGACAAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 720
QY 903 TTTTACTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
DB 721 TTTTACTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 1022
DB 781 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 840
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACCTT 1082

Db 841 GAAAGCGGTATGATCCCTTTGATCGCAGTCATCTTGAAACATGTTTCAACATCAAAATACGTT 900
Qy 1083 GATGTCGATACCAACGAATTTGCTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 1142
Db 901 GATGTCGATACCAACGAATTTGCTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 960
Qy 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTACTCTTACAAACATCTC 1202
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTTACTCTTACAAACATCTC 1020
Qy 1203 GATGCTTTTGGTATTATGGAATCTTAACTTAACTGGAAGATGAGGATTAATCAGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGGAATCTTAACTTAACTGGAAGATGAGGATTAATCAGATGAC 1080
Qy 1263 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1322
Db 1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1140
Qy 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCCTATGAT 1155

RESULT 3
I05204
LOCUS 1401 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 5 from Patent EP 0248227.
ACCESSION I05204
VERSION I05204.1 GI:591209
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1401)
AUTHORS Hagenson, M.J. and Stroman, D.W.
TITLE Yeast production of streptokinase
JOURNAL Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES
source 1..1401
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 69.2%; Score 1148.6; DB 6; Length 1401;
Best Local Similarity 99.2%; Pred. No. 1.9e-291;
Matches 1154; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 175 ATACCATGATTTGTCGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCCAAT 234
Db 2 ATTCATGATTTGTCGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCCAAT 61
Qy 235 TGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTT 294
Db 62 TAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTT 121
Qy 295 TTGAAATCGATCTTAACATCAGACCTGCTCATGGAGGAAGACAGACCAAGGCTTAAGTC 354
Db 122 TTGAAATCGATCTTAACATCAGACCTGCTCATGGAGGAAGACAGACCAAGGCTTAAGTC 181
Qy 355 CAAATCAAAACCATTTGCTAGTGTAGTGGCGCGATGTCACATAAACTTTGAGAAAGCTG 414
Db 182 CAAATCAAAACCATTTGCTAGTGTAGTGGCGCGATGTCACATAAACTTTGAGAAAGCTG 241
Qy 415 ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAACGACGACTACT 474
Db 242 ACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAACGACGACTACT 301
Qy 475 TTGAGTCAATTGATTTGCAAGCGATGCAACCATTTCTGATCGAAACGGCAAGGCTTACT 534
Db 302 TTGAGTCAATTGATTTGCAAGCGATGCAACCATTTCTGATCGAAACGGCAAGGCTTACT 361
Qy 535 TTGCTGACAAAGATGGTTCGGTAACTTCCGACCCCAACTGTCCAGAAATTTTGTCTAA 594

Db 362 TTGCTGACAAAGATGGTTTGGTAAACCTTTGCCGACCAACCTGTCCAGAAATTTTGTCTAA 421
Qy 595 GCGGACATGTGCGGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTG 654
Db 422 GCGGACATGTGCGGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTG 481
Qy 655 TTGATGTGGAATATATCTGTACAGATTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCCAG 714
Db 482 TTGATGTGGAATATATCTGTACAGATTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCCAG 541
Qy 715 GTCTCAAGATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCAATCAATCTCAAG 774
Db 542 GTCTCAAGATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCAATCAATCTCAAG 601
Qy 775 AATTACTAGCTCAAGCACAAGACATTTTAAACAAAAACCCAGGCTATACGATTTATG 834
Db 602 AATTACTAGCTCAAGCACAAGACATTTTAAACAAAAACCCAGGCTATACGATTTATG 661
Qy 835 AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 894
Db 662 AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGG 721
Qy 895 ATCAAGAGTTTACTTACCGTGTAAACATCGGGAACAAGCTTATAGGATCAATAAAAAAT 954
Db 722 ATCAAGAGTTTACTTACCGTGTAAACATCGGGAACAAGCTTATAGGATCAATAAAAAAT 781
Qy 955 CTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 1014
Db 782 CTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTA 841
Qy 1015 AAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACATGTTTCAACATCA 1074
Db 842 AAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACATGTTTCAACATCA 901
Qy 1075 AATACGTTGATGTGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGCG 1134
Db 902 AATACGTTGATGTGATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGCG 961
Qy 1135 AACGTAATGATGATTCAGAGATTTATACGATCTCTGATGATAGGCTAACTACTCTACA 1194
Db 962 AACGTAATGATGATTCAGAGATTTATACGATCTCTGATGATAGGCTAACTACTCTACA 1021
Qy 1195 ACATCTCGATGCTTTTGGTATTATGACTACTACTTAACTTAACTGGAAGCTAGAGGATATC 1254
Db 1022 ACAATCTCGATGCTTTTGGTATTATGACTACTACTTAACTTAACTGGAAGCTAGAGGATAATC 1081
Qy 1255 ACAGTACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATG 1314
Db 1082 ACAGTACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATG 1141
Qy 1315 CTAGCTATCATTTAGCCGGTGGT 1337
Db 1142 TCAGCTATCATTTAGCCCTATGAT 1164

RESULT 4
STRSK Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
LOCUS 2568 bp DNA linear BCT 26-APR-1993
DEFINITION Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
ACCESSION K02986
VERSION K02986.1 GI:153808
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
REFERENCE 1 (bases 1 to 2568)
AUTHORS Malke, H., Roe, B. and Ferretti, J. J.
TITLE Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A
JOURNAL Gene 34 (2-3), 357-362 (1985)
PUBMED 2389113
COMMENT Original source text: S.equisimilis (strain H46A) DNA, clone pMF5.

Draft entry and hard copy of sequence for [1] kindly provided by J.J. Ferretti, 03-SEP-1985.
The -35 and -10 regions are located at positions 760-765 and 781-786 respectively and an SD sequence at 808-813. Downstream from the coding region inverted repeats (positions 2176-2190 and 2203-2217) are thought to function as transcription terminators. The nucleotide sequence of skc does not support the hypothesis that the gene has evolved by duplication and fusion, as suggested by internal two-fold AA homologies of its product.

FEATURES

Location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
794..>2141
/product="skc mRNA"
819..2141
/note="prestreptokinase"
/codon_start=1
/transl_table=11
/protein_id="AAA26974.1"
/db_xref="GI:153809"
/translation="MKNYLSFGMFALLFALTFTGVNSVOAIGPEWLDRSPVNSQL
VYVAGTVEGTNQDLSLKFEDLTSRPAHGKTEQGLSPKSPKFPATDSGAMSHKLEK
ADLLKAIQBLJANYHSNDYFVIDFASDATTIDRQKVYADKDGSVTLTPQVQE
FLLSGHVRVPYKEPIONQAKSVDEVYVQTPPLNPDQDFRPLKDTLLKTLAIGD
TITSQELLAQAQSLINKHPGYTIYERDSSIYVHNDIIFRILPMQDEYFRVKNRQ
AIRINKSGINEEINLDLISIKYIVLKGEKPYDPDRSHLKLFTIKYVDVDTNELL
KSEQLLTASERNLDFRDLVDRKAKLYNLNDLAFGIMDYTLTKGVKEDHNDTNRIT
VTMGKPEGENASVHLAYDKRYTEEREVYSVLYRTGTPIPDNPNDK"
819..896
sig_peptide
/note="streptokinase signal peptide"
mat_peptide
897..2138
/product="streptokinase"
ORIGIN 5 bp upstream of PstI site.
Query Match 69.2%; Score 1148.6; DB 1; Length 2568;
Best Local Similarity 99.7%; Pred. No. 1.9e-291; Mismatches 4; Indels 0; Gaps 0;
Matches 1151; Conservative 0;
Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAATTTGGTTGTT 242
Db 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAATTTAGTTGTT 956
Qy 243 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 302
Db 957 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 1016
Qy 303 GATCTAAACATCAGACCTCTCATGGAGGAAGACAGACGAGGCTTAAGTCCAAATCA 362
Db 1017 GATCTAAACATCAGACCTCTCATGGAGGAAGACAGACGAGGCTTAAGTCCAAATCA 1076
Qy 363 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 1077 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 1136
Qy 423 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAAACGACACTTCTTGGAGTC 482
Db 1137 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAAACGACACTTCTTGGAGTC 1196
Qy 483 ATTGATTTTGCAGCGATCGAACCATTAATCTGATCGAAACGGCAAGCTCTACTTTGCTGAC 542
Db 1197 ATTGATTTTGCAGCGATCGAACCATTAATCTGATCGAAACGGCAAGCTCTACTTTGCTGAC 1256
Qy 543 AAAGATGTTTCGGTAACCTTTCGCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 602
Db 1257 AAAGATGTTTCGGTAACCTTTCGCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 1316
Qy 603 GTGCGCGTTAGACCATATATAAGAAAAAACAATACAAAAAACAAGCAAAATCTGTTGATGTG 662
Db 1317 GTGCGCGTTAGACCATATATAAGAAAAAACAATACAAAAAACAAGCAAAATCTGTTGATGTG 1376
Qy 663 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722

Db 1377 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1436
Qy 723 GATACTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 1437 GATACTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1496
Qy 783 GCTCAAGCACAAGACATTTTAAACAAAACCCACGAGCTATACGATTTTATGAACGTCAC 842
Db 1497 GCTCAAGCACAAGACATTTTAAACAAAACCCACGAGCTATACGATTTTATGAACGTCAC 1556
Qy 843 TCCTCAATCGTCACTCATGACAATTTTCGGTACGATTTTACCAATGGATCAAGAG 902
Db 1557 TCCTCAATCGTCACTCATGACAATTTTCGGTACGATTTTACCAATGGATCAAGAG 1616
Qy 903 TTTTACTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 1617 TTTTACTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1676
Qy 963 AATGAAGAAATAAACACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 1677 AATGAAGAAATAAACACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1736
Qy 1023 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAAATAGCTT 1082
Db 1737 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAACTGTTTCAACATCAAAATAGCTT 1796
Qy 1083 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 1797 GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1856
Qy 1143 TTAGACTTCAGAGATTTATACGATCTCTGATAGAGGCTAACTACTCTCAACAATCTC 1202
Db 1857 TTAGACTTCAGAGATTTATACGATCTCTGATAGAGGCTAACTACTCTCAACAATCTC 1916
Qy 1203 GATGCTTTTGGTATTATGAGTATACCTTAACTTGAAGGCTAGAGGATAATCAGCATGAC 1262
Db 1917 GATGCTTTTGGTATTATGAGTATACCTTAACTTGAAGGCTAGAGGATAATCAGCATGAC 1976
Qy 1263 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1322
Db 1977 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 2036
Qy 1323 CATTTAGCCGGTGGT 1337
Db 2037 CATTTAGCCGTATGAT 2051
RESULT 5
A04926
LOCUS A04926 2568 bp DNA linear PAT 15-JUL-1993
DEFINITION S.equisimilis skc gene for streptokinase.
ACCESSION A04926
VERSION A04926.1 GI:412219
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2568)
AUTHORS Ferretti, J.J. and Malke, H.
TITLE Streptokinase-coding recombinant vectors
JOURNAL Patent: EP 0151337-A 1 14-AUG-1985;
THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
FEATURES
source
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"

Query Match

69.2%; Score 1148.6; DB 6; Length 2568;

Best Local Similarity 99.7%; Pred. No. 1.9e-291; Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	183	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGCTGTTT	242
Db	897	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGCTGTT	956
QY	243	AGCGTTGCTGGTACTGTTGAGGGAGCGAATCAAGACATTAGTCTTAAATTTTTTGAATC	302
Db	957	AGCGTTGCTGGTACTGTTGAGGGAGCGAATCAAGACATTAGTCTTAAATTTTTTGAATC	1016
QY	303	GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACGAGCTTAAAGTCCAAATCA	362
Db	1017	GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACGAGCTTAAAGTCCAAATCA	1076
QY	363	AAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAATTTGAGAAGCTGACTTACTA	422
Db	1077	AAACCAATTTGCTACTGATAGTGGCGGATGTCACATAAATTTGAGAAGCTGACTTACTA	1136
QY	423	AAGGCTATTCAAGAACAAATTTGATCGCTTAAACGTCACAGTAAACGACGACTACTTTGAGTC	482
Db	1137	AAGGCTATTCAAGAACAAATTTGATCGCTTAAACGTCACAGTAAACGACGACTACTTTGAGTC	1196
QY	483	ATTGATTTTGAAAGCGATGCAACCAATTAATCTGATCGAAACGCAAGGCTCTACTTTGCTGAC	542
Db	1197	ATTGATTTTGAAAGCGATGCAACCAATTAATCTGATCGAAACGCAAGGCTCTACTTTGCTGAC	1256
QY	543	AAAGATGTTTCGGTAACTTCGCGACCCACCTGTCGAAGAAATTTTCTAAGCGGCAT	602
Db	1257	AAAGATGTTTCGGTAACTTCGCGACCCACCTGTCGAAGAAATTTTCTAAGCGGCAT	1316
QY	603	GTGCGGTTAGACATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	662
Db	1317	GTGCGGTTAGACATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	1376
QY	663	GAATFACTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA	722
Db	1377	GAATFACTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA	1436
QY	723	GATCTAAGCTATTGAAACACTAGTATCGGTGACCAATCATCTCAAGAAATTA	782
Db	1437	GATCTAAGCTATTGAAACACTAGTATCGGTGACCAATCATCTCAAGAAATTA	1496
QY	783	GCTCAAGCAAAAGCATTTTAAACAAAAACCCACGAGCTATACGTTTATGACGTC	842
Db	1497	GCTCAAGCAAAAGCATTTTAAACAAAAACCCACGAGCTATACGTTTATGACGTC	1556
QY	843	TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	902
Db	1557	TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	1616
QY	903	TTTACTTACCGTTTAAAAATCGGGAACAGCTTATAGATCAATAAAAAATCTGGTCTG	962
Db	1617	TTTACTTACCGTTTAAAAATCGGGAACAGCTTATAGATCAATAAAAAATCTGGTCTG	1676
QY	963	AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	1022
Db	1677	AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	1736
QY	1023	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATAGCTT	1082
Db	1737	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATAGCTT	1796
QY	1083	GATGTCGATCAACAGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGTACGCAACGTAAC	1142
Db	1797	GATGTCGATCAACAGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGTACGCAACGTAAC	1856
QY	1143	TTAGACTTCAGAGATTTATACGATCTCTCGTGATAGGCTTAACTACTCTACACAATCTC	1202
Db	1857	TTAGACTTCAGAGATTTATACGATCTCTCGTGATAGGCTTAACTACTCTACACAATCTC	1916
QY	1203	GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTACAGGATATACGATGAC	1262

Db	1917	GATGCTTTGGTATTATGAGCTATATCTTAACTGAAAGTAGAGATAATCACGATGAC	1976
QY	1263	ACCAACCGTATCATCAACCGTTTATATGGCAAGCCAGCCGAGGAGAGATCTAGCTAT	1322
Db	1977	ACCAACCGTATCATCAACCGTTTATATGGCAAGCCAGCCGAGGAGAGATCTAGCTAT	2036
QY	1323	CATTTAGCCGCTGGT	1337
Db	2037	CATTTAGCCCTATGAT	2051
RESULT 6			
SEDEXB			
LOCUS	SEDEXB	8931 bp	DNA linear BCT 18-APR-2005
DEFINITION	S.equisimilis dextb, abc, lrp, skc, rel genes and ORF1.		
ACCESSION	X72832		
VERSION	X72832.1	GI:407876	
KEYWORDS	abc gene; ABC transporter; dexB gene; dextran glucosidase; glucan 1,6-alpha-glucosidase; leucine rich protein; LRP gene; ORF1; rel gene; skc gene; streptokinase; stringent response-like protein. Streptococcus dysgalactiae subsp. equisimilis Streptococcus dysgalactiae subsp. equisimilis Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
SOURCE			
ORGANISM			
REFERENCE	1 (bases 3621 to 6190)		
AUTHORS	Malke,H., Roe,B. and Ferretti,J.J		
TITLE	Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A		
JOURNAL	Gene 34 (2-3), 357-362 (1985)		
PUBMED	298113		
REFERENCE	2 (bases 1 to 4188; 5790 to 8931)		
AUTHORS	Mechold,U., Steiner,K., Vettermann,S. and Malke,H.		
TITLE	Genetic organization of the streptokinase region of the Streptococcus equisimilis H46A chromosome		
JOURNAL	Mol. Gen. Genet. 241 (1-2), 129-140 (1993)		
PUBMED	8232196		
REFERENCE	3 (bases 1 to 8931)		
AUTHORS	Malke,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-MAY-1993) H. Malke, Institute for Molecular Biology, Jena University, Winzerlaer Str 10, 07708 Jena, FRG		
COMMENT	Related sequences: K02986, M19346, X13399 & X13400.		
FEATURES			
source	1..8931		
	/organism="Streptococcus dysgalactiae subsp. equisimilis"		
	/mol_type="genomic DNA"		
	/strain="H46A"		
	/isolate="human group C strain"		
	/sub_species="equisimilis"		
	/db_xref="taxon:119602"		
	/chromosome="streptokinase region"		
	/clone_lib="E.coli plasmid library containing subfragments of the submitted seq:PSHD14/16, pSPV19, pSH2, pMP1, pCWP73, pRH10, pWX4"		
	complement(89..1761)		
gene	/gene="dexB"		
	complement(89..115)		
terminator	/gene="dexB"		
	/note="hairpin loop"		
	complement(136..1749)		
CDS	/gene="dexB"		
	/EC number="3.2.1.70"		
	/codon_start=1		
	/transl_table=11		
	/product="glucan 1,6-alpha-glucosidase"		
	/protein_id="CAA51348.1"		
	/db_xref="GI:407877"		
	/db_xref="GOA:Q59905"		
	/db_xref="InterPro:IPR006047"		
	/db_xref="UniProt:IPR006589"		
	/db_xref="UniProt/Swiss-Prot:O59905"		
	/translation="MQQWVKATIIYIYPRSPKDTSGNGIGDLKIGTSDLYLQKLG ITAIWLSPVYQSPMDNGYDISDYEAIEAFVGNMDDMLLAAANERGIKIIMDLVNN		

Db 4517 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACACCAATAGTTGTT 4576
Qy 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 4577 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 4636
Qy 303 GATCTAAATCATCAGCACTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 362
Db 4637 GATCTAAATCATCAGCACTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 4696
Qy 363 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTTACTA 422
Db 4697 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTTACTA 4756
Qy 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAAACGACTACTTTGAGGTC 482
Db 4757 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAAACGACTACTTTGAGGTC 4816
Qy 483 ATTGATTTTGAAGCGATGCAACCAATTACTGTATCGAAACGGCAAGGTCTACTTTGCTGAC 542
Db 4817 ATTGATTTTGAAGCGATGCAACCAATTACTGTATCGAAACGGCAAGGTCTACTTTGCTGAC 4876
Qy 543 AAAGATGGTTCGGTAACCTTTGCCGACCCAACTGTCCAGAAATTTTGTCTAAGCGGCAT 602
Db 4877 AAAGATGGTTCGGTAACCTTTGCCGACCCAACTGTCCAGAAATTTTGTCTAAGCGGCAT 4936
Qy 603 GTGCGGTTAGACCATATAAGAAACCAATCAACCAACCAAGCGAATCTGTTGATGTG 662
Db 4937 GTGCGGTTAGACCATATAAGAAACCAATCAACCAACCAAGCGAATCTGTTGATGTG 4996
Qy 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAACAGAGTCTCAAA 722
Db 4997 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAACAGAGTCTCAAA 5056
Qy 723 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 5057 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 5116
Qy 783 GCTCAAGCAAAAGCAATTTAAACAAACCAACCCAGGCTATAGGATTTATGAACGTGAC 842
Db 5117 GCTCAAGCAAAAGCAATTTAAACAAACCAACCCAGGCTATAGGATTTATGAACGTGAC 5176
Qy 843 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 902
Db 5177 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 5236
Qy 903 TTTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 5237 TTTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 5296
Qy 963 AATGAAGAAATAAACCAACTGACCTGATCTTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 5297 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 5356
Qy 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 1082
Db 5357 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 5416
Qy 1083 GATGTCGATACCAACCAATTTGCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 5417 GATGTCGATACCAACCAATTTGCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 5476
Qy 1143 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGGCTAAACTACTCTCAACAATCTC 1202
Db 5477 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGGCTAAACTACTCTCAACAATCTC 5536
Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTTGAAAAAGTAGAGGATAATCAGATGAC 1262
Db 5537 GATGCTTTTGGTATTATGACTATACCTTAACCTTGAAAAAGTAGAGGATAATCAGATGAC 5596
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGCAAGGAGAGTCTAGCTAT 1322

Db 5597 ACCAACCGTATCATACCGTTTATATGGGCAAGCCGACCCGAGAGAGATGCTAGCTAT 5656
Qy 1323 CATTTAGCCGGTGGT 1337
Db 5657 CATTTAGCCTATGAT 5671
RESULT 7
E00522
LOCUS E00522
DEFINITION DNA fragment of plasmid PMF1 into which DNA encoding streptokinase is inserted.
ACCESSION E00522
VERSION E00522.1 GI:2168801
KEYWORDS JP 1985237995-A/1.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 2568)
AUTHORS Jiyosefu,J.F. and Horusuto,M.
TITLE RECOMINATION VECTOR FOR PRODUCING STREPTOKINASE
JOURNAL Patent: JP 1985237995-A 1 26-NOV-1985;
PHILLIPS PETROLEUM CO
COMMENT OS Streptococcus equisimilis
PN JP 1985237995-A/1
PD 26-NOV-1985
PF 09-OCT-1984 JP 1984212403
PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PI
PJ JIYOSFU JIEI FUERETSUTEI, HORUSUTO MARUKE
PC C12N15/00, C12N1/20, C12N9/70, (C12N15/00, C12R1:46), (C12N1/20, PC C12R1:19),
PC (C12N9/70, C12R1:19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Streptococcus equisimilis H46A; CC *source: library=streptococcus equisimilis H46 library; CC *source: clone=lambda L47 skc clone;
FH Key Location/Qualifiers
FT 5'UTR 1..818
FT sig_peptide 819..896 /product='streptokinase signal peptide' FT
FT mat_peptide 897..2138 /product='streptokinase'
FT CDS 819..2143 /product='streptokinase'
FT 3'UTR 2142..2569.
FEATURES
source Location/Qualifiers
1..2568
/organism='Streptococcus dysgalactiae subsp. equisimilis'
/mol_type='genomic DNA'
/db_xref='taxon:119602'
ORIGIN
Query Match 68.9%; Score 1143.8; DB 6; Length 2568;
Best Local Similarity 99.4%; Pred. No. 3.6e-290;
Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTTGTT 242
Db 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTAGTTGTT 956
Qy 243 AGCGTTCGTCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 957 AGCGTTCGTCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 1016
Qy 303 GATCTAAATCATCAGCACTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 362
Db 1017 GATCTAAATCATCAGCACTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAAAATCA 1076

QY 363 AAACCAATTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
 Db |||||
 QY 1077 AAACCAATTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 1136
 Db |||||
 QY 423 AAGGCTATTCAAGAACAAATTGATCGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 482
 Db |||||
 QY 1137 AAGGCTATTCAAGAACAAATTGATCGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 1196
 Db |||||
 QY 483 ATTGATTTTGAAGGATCAACCAATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
 Db |||||
 QY 1197 ATTGATTTTGAAGGATCAACCAATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 1256
 Db |||||
 QY 543 AAAGATGTTTCCGCTAACCTTGGCGACCCCAACCTGTCCAAGAAATTTTGTCTAAACGGACAT 602
 Db |||||
 QY 1257 AAAGATGTTTCCGCTAACCTTGGCGACCCCAACCTGTCCAAGAAATTTTGTCTAAACGGACAT 1316
 Db |||||
 QY 603 GTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCAAACTGTGTGATGTG 662
 Db |||||
 QY 1317 GTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGCAAACTGTGTGATGTG 1376
 Db |||||
 QY 663 GAAATACGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTGACACGAGTCTCAAA 722
 Db |||||
 QY 1377 GAAATACGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTGACACGAGTCTCAAA 1436
 Db |||||
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 782
 Db |||||
 QY 1437 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 1496
 Db |||||
 QY 783 GCTCAAGCAAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db |||||
 QY 1497 GCTCAAGCAAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1556
 Db |||||
 QY 843 TCCTCAAGCTCACTCATGACAAATGACATTTTCGCTAGCAATTTTACCAATGGATCAAGAG 902
 Db |||||
 QY 1557 TCCTCAAGCTCACTCATGACAAATGACATTTTCGCTAGCAATTTTACCAATGGATCAAGAG 1616
 Db |||||
 QY 903 TTTACTACCGTGTAAACCTCGGAAACCAAGCTTATAGGATCAATAAAACCTGTGCTG 962
 Db |||||
 QY 1617 TTTACTACCGTGTAAACCTCGGAAACCAAGCTTATAGGATCAATAAAACCTGTGCTG 1676
 Db |||||
 QY 963 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db |||||
 QY 1677 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1736
 Db |||||
 QY 1023 GAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAAACTGTTCACATCAAAATCGTT 1082
 Db |||||
 QY 1737 GAAAGCGGTATGATCCCTTTGATGCGAGTCACTTGAAACTGTTCACATCAAAATCGTT 1796
 Db |||||
 QY 1083 GATGTCGATACCAACGAAATTTGCTTAAAGTGAAGTCACTTAAACGCTAGCGAAGCTAAC 1142
 Db |||||
 QY 1797 GATGTCGATACCAACGAAATTTGCTTAAAGTGAAGTCACTTAAACGCTAGCGAAGCTAAC 1856
 Db |||||
 QY 1143 TTAGACTTTCAGAGATTTATACGATCTCTGATTAAGCTTAAACCTTCTTACAAACATCTC 1202
 Db |||||
 QY 1857 TTAGACTTTCAGAGATTTATACGATCTCTGATTAAGCTTAAACCTTCTTACAAACATCTC 1916
 Db |||||
 QY 1203 GATGCTTTTGGTATTATGACATTAACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1262
 Db |||||
 QY 1917 GATGCTTTTGGTATTATGACATTAACCTTAACTGGAAGAGTAGAGGATATCAAGTAC 1976
 Db |||||
 QY 1263 ACCAACCGTATCAACCGTTTATATGGCGAAGCGACCGGAGGAGAGATGTAGCTAT 1322
 Db |||||
 QY 1977 ACCAACCGTATCAACCGTTTATATGGCGAAGCGACCGGAGGAGAGATGTAGCTAT 2036
 Db |||||
 QY 1323 CATTTAGCGGTTGGT 1337
 Db |||||
 QY 2037 CATTTAGCGTATGAT 2051
 Db |||||

RESULT 8
 QY97820 7057 bp DNA linear PAT 20-APR-2004
 LOCUS
 DEFINITION Sequence 26 from Patent WO2004029256.

ACCESSION CQ797820
 VERSION CQ797820.1 GI:46426093
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 other sequences; artificial sequences.
 REFERENCE 1
 AUTHORS selman-housen Sosa,G., aguiar Cabeza,E., gonz lez Quintero,A.D.
 and ramos gonz Lez,O.
 Vector for the production of transplastomic angiosperm plants
 Patent: WO 2004029256-A 26 08-APR-2004;
 JOURNAL CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
 FEATURES
 Location/Qualifiers
 source 1..7057
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Description of Artificial Sequence: Nucleotide
 sequence of DNA fragment from the vector pVTPA-Estrep
 between the rice atpB and tobacco rbcL borders."
 ORIGIN
 Query Match 68.4%; Score 1136.2; DB 6; Length 7057;
 Best Local Similarity 98.5%; Pred. No. 3.8e-288;
 Matches 1147; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 QY 173 ATATACCATGATTCGTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCA 232
 Db |||||
 QY 2013 ATAAACCATGATTCGTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCA 2072
 Db |||||
 QY 233 ATTGGTTGTTAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATT 292
 Db |||||
 QY 2073 ATTAGTTGTTAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATT 2132
 Db |||||
 QY 293 TTTTGAATCGATCTAAACATCACGACCTGCTCATGAGGAGAAAGACAGACGAAGCTTAAAG 352
 Db |||||
 QY 2133 TTTTGAATTTGACCTTAACATCACGACCTGCTCATGAGGAGAAAGACAGACGAAGCTTAAAG 2192
 Db |||||
 QY 353 TCACAAATCAAAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTTGAGAAAGC 412
 Db |||||
 QY 2193 TCCAAATCAAAACCATTTGCTACTGATAGTGGCGGATGTCCATAACTTTGAAAAAGC 2252
 Db |||||
 QY 413 TGACTTACTAAAGCGTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTA 472
 Db |||||
 QY 2253 TGACTTACTAAAGCGTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTA 2312
 Db |||||
 QY 473 CTTTGAGGTCATTGATTTTTCAGACGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTA 532
 Db |||||
 QY 2313 CTTTGAGGTCATTGATTTTTCAGACGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTA 2372
 Db |||||
 QY 533 CTTTGCTGACAAAGATGGTTGCGTAAACCTTGGCGGACCCCAACCTGTCCAAGAAATTTTGCT 592
 Db |||||
 QY 2373 CTTTGCTGACAAAGATGGTTGCGTAAACCTTGGCGGACCCCAACCTGTCCAAGAAATTTTGCT 2432
 Db |||||
 QY 593 AAGCGGACATGTCGCGTTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATC 652
 Db |||||
 QY 2433 AAGCGGACATGTCGCGTTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATC 2492
 Db |||||
 QY 653 TGTGTGATGTCGAATATATCTGTACAGTTTACTCTCCCTTAAACCTGATGACGATTTTCAGACC 712
 Db |||||
 QY 2493 TGTGTGATGTCGAATATATCTGTACAGTTTACTCTCCCTTAAACCTGATGACGATTTTCAGACC 2552
 Db |||||
 QY 713 AGGTCTCAAAAGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCA 772
 Db |||||
 QY 2553 AGGTCTCAAAAGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCA 2612
 Db |||||
 QY 773 AGAATTAAGTCTCAAGCACAAAGCATTTTAAACAAAAACCCACCGGCTATACGATTTA 832
 Db |||||
 QY 2613 AGAATTAAGTCTCAAGCACAAAGCATTTTAAACAAAAACCCACCGGCTATACGATTTA 2672
 Db |||||
 QY 833 TGAACGTGACTCTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAGT 892
 Db |||||
 QY 2673 TGAACGTGACTCTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAGT 2732
 Db |||||

QY 893 GGATCAAGAGTTTACTTACCGTGTATAAATCGGAAACAAGCTTATAGGATCAATAAAAA 952
|||||
Db 2733 GGATCAAGAGTTTACTTACCATGTCAAAATCGGAAACAAGCTTATGAGATCAATAAAAA 2792
QY 953 ATCTGCTCTGAATGAAGAATAAACACACAGCTGATCTCTGAGAAATATTACGTCCT 1012
|||||
Db 2793 ATCTGCTCTGAATGAAGAATAAACACACAGCTGATCTCTGAGAAATATTACGTCCT 2852
QY 1013 TAAAAAGGGAAAAAGCCGTATGATCTCTTTGATCGCAGTCACTTTGAAACTGTTTACCAT 1072
|||||
Db 2853 TAAAAAGGGAAAAAGCCGTATGATCTCTTTGATCGCAGTCACTTTGAAACTGTTTACCAT 2912
QY 1073 CAAATACGTTGATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAG 1132
|||||
Db 2913 CAAATACGTTGATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAG 2972
QY 1133 CGAACCTTAACCTTAGACTTACGAGATTTATACGATCTCTCGTGATAGAGCTTAACTACTTA 1192
|||||
Db 2973 CGAACCTTAACCTTAGACTTACGAGATTTATACGATCTCTCGTGATAGAGCTTAACTACTTA 3032
QY 1193 CAAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAA 1252
|||||
Db 3033 CAAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAA 3092
QY 1253 TCAAGATGACACCAACCGCTATCATACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAA 1312
|||||
Db 3093 TCAAGATGACACCAACCGCTATCATACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAA 3152
QY 1313 TGCTAGCTATCATATTAGCCGCTGGT 1337
|||||
Db 3153 TGCTAGCTATCATATTAGCCCTATGAT 3177

RESULT 9
AR068768
LOCUS
DEFINITION
Sequence 19 from patent US 5854049.
ACCESSION
AR068768
VERSION
AR068768.1 GI:6000975
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 2566)
AUTHORS
Reed,G.L.
TITLE
Plasmin-resistant streptokinase
JOURNAL
Patent: US 5854049-A 19 29-DEC-1998;
FEATURES
Location/Qualifiers
source
1..2566
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 68.3%; Score 1134.6; DB 6; Length 2566;
Best Local Similarity 99.5%; Pred. No. 9.5e-288;
Matches 1149; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGGTTGT 242
|||||
Db 897 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTAGTTGT 956
QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 302
|||||
Db 957 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 1016
QY 303 GATCTAAACATCAGACCTCTCATGAGGAGAGACAGACGACGCTTAAGTCCAAATCA 362
|||||
Db 1017 GATCTAAACATCAGACCTCTCAT--AGGAAGACAGACGACGCTTAAGTCCAAATCA 1074
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTACAGAAAGCTGACTTACTA 422
|||||
Db 1075 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTACAGAAAGCTGACTTACTA 1134

QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCCACAGTAACGACGACTACTTTTGGGTC 482
|||||
Db 1135 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCCACAGTAACGACGACTACTTTTGGGTC 1194
QY 483 ATTGATTTTGGCAAGGATGCAACCAATTAATGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
|||||
Db 1195 ATTGATTTTGGCAAGGATGCAACCAATTAATGATCGAAACGGCAAGGTCTACTTTGCTGAC 1254
QY 543 AAAGATGGTTCGGTAAACCTTCCGACCCCAACCTGTCTCCAAAGATTTTGTCTAAGCCGACAT 602
|||||
Db 1255 AAAGATGGTTCGGTAAACCTTCCGACCCCAACCTGTCTCCAAAGATTTTGTCTAAGCCGACAT 1314
QY 603 GTGCGCGTTTACACCATATAAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTCATGTG 662
|||||
Db 1315 GTGCGCGTTTACACCATATAAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTCATGTG 1374
QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGTCTCAAA 722
|||||
Db 1375 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGTCTCAAA 1434
QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 782
|||||
Db 1435 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 1494
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 842
|||||
Db 1495 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1554
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTTCGTACGATTTTACCAATGGATCAAGAG 902
|||||
Db 1555 TCCTCAATCGTCACTCATGACAAATGACATTTTTCGTACGATTTTACCAATGGATCAAGAG 1614
QY 903 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
|||||
Db 1615 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1674
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGTCTTTAAAAAAGG 1022
|||||
Db 1675 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGTCTTTAAAAAAGG 1734
QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACCATCAAAATACGT 1082
|||||
Db 1735 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACCATCAAAATACGT 1794
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAACAGCTTAGCGAAACGTAAC 1142
|||||
Db 1795 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAACAGCTTAGCGAAACGTAAC 1854
QY 1143 TTAGACTTCAGAGATTTTATACGATCTCTGATAGGCTAAACCTACTCTACAAATCTC 1202
|||||
Db 1855 TTAGACTTCAGAGATTTTATACGATCTCTGATAGGCTAAACCTACTCTACAAATCTC 1914
QY 1203 GATGCTTTTGGTATTATAGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
|||||
Db 1915 GATGCTTTTGGTATTATAGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1974
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1322
|||||
Db 1975 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 2034
QY 1323 CATTTAGCCGCTGGT 1337
|||||
Db 2035 CATTTAGCCCTATGAT 2049

RESULT 10
S46536
LOCUS
DEFINITION
SKC-2=streptokinase [Streptococcus equisimilis, group C, ATCC 9542,
Genomic, 1245 nt].
ACCESSION
S46536
VERSION
S46536.1 GI:257196

KEYWORDS

SOURCE Streptococcus dysgalactiae subsp. equisimilis
 ORGANISM Streptococcus dysgalactiae subsp. equisimilis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.

REFERENCE

AUTHORS 1 (bases 1 to 1245)
 Estrada,M.P., Hernandez,L., Perez,A., Rodriguez,P., Serrano,R.,
 Rubiera,R., Pedraza,A., Padron,G., Antuch,W., de la Fuente,J. et
 al.

TITLE High level expression of streptokinase in Escherichia coli

JOURNAL Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)

PUBMED 1368792

REMARK GenBank staff at the National Library of Medicine created this
 entry [NCBI gibbsq 115306] from the original journal article.

FEATURES

source
 1..1245
 /organism="Streptococcus dysgalactiae subsp. equisimilis"
 /mol_type="genomic DNA"
 /db_xref="taxon:119602"
 /note="group: C"

gene

1..1245
 /gene="SKC-2"
 /protein_id="AAC60418.1"
 /db_xref="GI:257197"
 /translation="TAGPEWLLDRPSVNNLSQVSVAGTVGEGTNOISLKFPEDILTS
 /translat="GI:257197"
 RPAGGKTEOGLSPKSPKPPATDSGPHKLEKADLLKAIOELIANVHNSDNDYFEVID
 PASDATITIDRKNVYFADKDGSVTLPTQVQFELLSCHVRVPYKPIQNAQKSDV
 EYTVQTFELNPDPPDRPGLKDTLLKTLAIGDTITSOELLAQAQSLINKHPHYIYE
 RDSSIVHDNIDFRIPLPMDQFTVHVNRQAYEINKSGLEENETDLDLISEKYV
 LKKGKPYDPFDRSLKPLFTIKYVDVNTNELLSEQLLTASERNLDFRDLVPRDKAK
 LLYNLDFAFGIMDYTLTGKVEDNDHTNRITVYMGKREGENASVHLYADKDRYTEE
 EREVSYLRYTGTPIPDNPNDR"

CDS

1..1245
 /gene="SKC-2"
 /note="streptokinase; Mature protein lacks initial Met"
 /codon_start=1
 /product="SKC-2"
 /protein_id="AAC60418.1"
 /db_xref="GI:257197"
 /translation="TAGPEWLLDRPSVNNLSQVSVAGTVGEGTNOISLKFPEDILTS
 /translat="GI:257197"
 RPAGGKTEOGLSPKSPKPPATDSGPHKLEKADLLKAIOELIANVHNSDNDYFEVID
 PASDATITIDRKNVYFADKDGSVTLPTQVQFELLSCHVRVPYKPIQNAQKSDV
 EYTVQTFELNPDPPDRPGLKDTLLKTLAIGDTITSOELLAQAQSLINKHPHYIYE
 RDSSIVHDNIDFRIPLPMDQFTVHVNRQAYEINKSGLEENETDLDLISEKYV
 LKKGKPYDPFDRSLKPLFTIKYVDVNTNELLSEQLLTASERNLDFRDLVPRDKAK
 LLYNLDFAFGIMDYTLTGKVEDNDHTNRITVYMGKREGENASVHLYADKDRYTEE
 EREVSYLRYTGTPIPDNPNDR"

ORIGIN

Query Match 67.9%; Score 1127.8; DB 1; Length 1245;
 Best Local Similarity 98.5%; Pred. No. 5.7e-286;
 Matches 1138; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTTGGTTGT 242
 Db 1 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTTAGTTGT 60
 Qy 243 AGCGTTGCTGTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
 Db 61 AGCGTTGCTGTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 120
 Qy 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 362
 Db 121 GACCTAACATCAGCACTGCTCATGGAGGAAAGACAGACAGGCTTAAAGTCCAAATCA 180
 Qy 363 AAACCAATTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGCTACTA 422
 Db 181 AAACCAATTGCTACTGATAGTGGCGGATGCCACATAAACTTGAGAAAGCTGCTACTA 240
 Qy 423 AAGGCTATTCAAGAACAAATTGATCGGTAACTGCTCCAGTAAACGACACTACTTTGAGTC 482
 Db 241 AAGGCTATTCAAGAACAAATTGATCGGTAACTGCTCCAGTAAACGACACTACTTTGAGTC 300
 Qy 483 ATTGATTTTGAAGCGATCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
 Db 301 ATTGATTTTGAAGCGATCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 360
 Qy 543 AAAGATGGTTCGGTAACCTTGGCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
 Db 361 AAAGATGGTTCGGTAACCTTGGCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 420
 Qy 603 GTGCGCGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATGTC 662

Db 421 GTGCGGTTAGACCATATAAGAAACCAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
 Qy 663 GATATACGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 722
 Db 481 GAATATACGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAAA 540
 Qy 723 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 782
 Db 541 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
 Qy 783 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC 842
 Db 601 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC 660
 Qy 843 TCCTCAATCGTCACTCATGACCAATGACATTTTCGGTAGCATTTTACCAATGGATCAAGAG 902
 Db 661 TCCTCAATCGTCACTCATGACCAATGACATTTTCGGTAGCATTTTACCAATGGATCAAGAG 720
 Qy 903 TTTTACTTACCGTGTAAATAATCGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 721 TTTTACTTACCATGTCAAAAATCGGAAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 780
 Qy 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 Db 781 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
 Qy 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTTACCATCAATAGCTT 1082
 Db 841 GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTTACCATCAATAGCTT 900
 Qy 1083 GATGTGTCATACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 901 GATGTCAACACCAACGAATTTGCTAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
 Qy 1143 TTAGACTTCAGAGATTTTATACGATCCCTGCTGATAAGGCTAAACTACTCTACAACAATCTC 1202
 Db 961 TTAGACTTCAGAGATTTTATACGATCCCTGCTGATAAGGCTAAACTACTCTACAACAATCTC 1020
 Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATATACAGATGAC 1262
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATATACAGATGAC 1080
 Qy 1263 ACCAAGCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1322
 Db 1081 ACCAAGCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
 Qy 1323 CATTTAGCCGGTGGT 1337
 Db 1141 CATTTAGCCCTATGAT 1155

RESULT 11

A20015
 LOCUS 1257 bp DNA linear PAT 14-JUL-1995
 DEFINITION SEQ ID NO: 23; Nucleotide sequence for methionol-streptokinase
 fusion protein.
 ACCESSION A20015
 VERSION A20015.1 GI:1247848
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM synthetic construct
 other sequences; artificial sequences.
 REFERENCE 1 (bases 1 to 1257)
 AUTHORS PROTEINS AND NUCLEIC ACIDS
 TITLE Patent: WO 9109125-A 23 27-JUN-1991;
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1..1257
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 4..1251
 /codon_start=1

CDS


```
/transl_table=11
/product="methionyl-streptokinase fusion protein"
/protein_id="CAA01486.1"
/db_xref="GI:1247849"
/transl_table="1"
/translation="MAGPEWLLDRPSVNNNSQVSVAGVTEGNDQISLKFPEIDIT
SPAHGKTEQGLSPKSKPENTDSAMPKLEKADLLKAEQLIANVHNDITFEVI
DFASADITDRNGKVFPADKDSVLTQPVQBFLLSGHVRVPRPKPIONQAKSV
VRYTQFTPLNPDNDPFGKLDKTLKLAIGDTTISQELLAQAQSLNTHPGYTIY
ERDSSTVTHDNDIFRITLPMQDQFVHVKNREQAYEINKSGLNNEINNTDLISEKY
VLKSGEPYDPPDRSHLKLFTI KYVDVNTNELLKSEOLLTKSBRNLDPRLDYPDRKA
KLLYNLDAFGIMDVTLTGKVDENHDDTNRITIVYMKRPEGENASVHLAYDKDRYTE
EBREYVSYLRYGTGPIPDNDK"
```

ORIGIN

```
Query Match 67.9%; Score 1127.6; DB 6; Length 1257;
Best Local Similarity 98.4%; Pred. No. 6.5e-286;
Matches 1139; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 180 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGGTT 239
DB 4 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATAGTT 63

QY 240 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGA 299
DB 64 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGA 123

QY 300 ATCGATCTAACTACGACCTGCTGATGAGGAGAAAGACAGACAGAGCTTAAAGTCCAAA 359
DB 124 ATGACCTTAACATCAGACTGCTCATGAGGAAAGACAGACAGAGCTTAAAGTCCAAA 183

QY 360 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAACTTGGAGAAAGCTGACTTA 419
DB 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAACTTGGAGAAAGCTGACTTA 243

QY 420 CTAAGGCTATTCAAGAACAAATTTGATCGTCCACAGTACAGACGACTACTTTGAG 479
DB 244 CTAAGGCTATTCAAGAACAAATTTGATCGTCCACAGTACAGACGACTACTTTGAG 303

QY 480 GTCAATGATTTTGAAGCGATGCAACCATTAAGTATGATGCAACCGCAGGTCTACTTTGCT 539
DB 304 GTCAATGATTTTGAAGCGATGCAACCATTAAGTATGATGCAACCGCAGGTCTACTTTGCT 363

QY 540 GACAAAGATGGTTGGTAACTTGGTCCGACCAACCTGTCCAGAAATTTTGGTAAAGCGGA 599
DB 364 GACAAAGATGGTTGGTAACTTGGTCCGACCAACCTGTCCAGAAATTTTGGTAAAGCGGA 423

QY 600 CATGTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 659
DB 424 CATGTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGAT 483

QY 660 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTCAAGACAGGTCTC 719
DB 484 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTCAAGACAGGTCTC 543

QY 720 AAAGATACCTAAGCTATTGAAACACATAGTATCGGTGACACCATCAATCTCAAGAAATTA 779
DB 544 AAAGATACCTAAGCTATTGAAACACATAGTATCGGTGACACCATCAATCTCAAGAAATTA 603

QY 780 CTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCGGCTATACGATTTATGAACGT 839
DB 604 CTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCGGCTATACGATTTATGAACGT 663

QY 840 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCGTCAGATTTTACCAATGGATCAA 899
DB 664 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCGTCAGATTTTACCAATGGATCAA 723

QY 900 GAGTTTACTTACCGTTTAAATTCGGGACAGCTTATAGGATCAATAAAATCTGGT 959
DB 724 GAGTTTACTTACCGTTTAAATTCGGGACAGCTTATAGGATCAATAAAATCTGGT 783

QY 960 CTGAATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAA 1019
DB 784 CTGAATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAA 843
```

```
QY 1020 GGGAAAGCCGTATGATCCCTTTGATCCAGTCACTTGAACACTGTTCAACATCAATAC 1079
DB 844 GGGAAAGCCGTATGATCCCTTTGATCCAGTCACTTGAACACTGTTCAACATCAATAC 903

QY 1080 GTTGATGTCGATACCAACGAATTTGCTAAAGTGAAGAGCTTTAAACAGCTAGCGAAGT 1139
DB 904 GTTGATGTCGATACCAACGAATTTGCTAAAGTGAAGAGCTTTAAACAGCTAGCGAAGT 963

QY 1140 AACTTAGACTTTCAGAGATTTATACGATCCTCGTGAAGGCTTAAGTCTTACAACTAT 1199
DB 964 AACTTAGACTTTCAGAGATTTATACGATCCTCGTGAAGGCTTAAGTCTTACAACTAT 1023

QY 1200 CTCGATGCTTTTGGTATTATGACACTATACCTTTAACTGGAAGTAGAGATAATCACCAT 1259
DB 1024 CTCGATGCTTTTGGTATTATGACACTATACCTTTAACTGGAAGTAGAGATAATCACCAT 1083

QY 1260 GACACCAACCGTATCATAACCGTTTATATGGCAAGCGACCGAGAGAGAGATCTAGC 1319
DB 1084 GACACCAACCGTATCATAACCGTTTATATGGCAAGCGACCGAGAGAGAGATCTAGC 1143

QY 1320 TATCATTTAGCCGCTGGT 1337
DB 1144 TATCATTTAGCCCTATGAT 1161

RESULT 12
LOCUS I13203 1257 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 25 from patent US 5434073.
ACCESSION I13203
VERSION I13203.1 GI:910551
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1257)
AUTHORS Dawson, K., Hunter, M. G. and Czaplewski, L. G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 25 JUL-1995;
FEATURES Location/Qualifiers
source 1..1257
/organism="unknown"
/mol_type="unassigned DNA"
```

```
Query Match 67.9%; Score 1127.6; DB 6; Length 1257;
Best Local Similarity 98.4%; Pred. No. 6.5e-286;
Matches 1139; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 180 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATGGTT 239
DB 4 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATAGTT 63

QY 240 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGA 299
DB 64 GTTAGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGA 123

QY 300 ATCGATCTAACTACGACCTGCTGATGAGGAGAAAGACAGACAGAGCTTAAAGTCCAAA 359
DB 124 ATGACCTTAACATCAGACTGCTCATGAGGAAAGACAGACAGAGCTTAAAGTCCAAA 183

QY 360 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAACTTGGAGAAAGCTGACTTA 419
DB 184 TCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAACTTGGAGAAAGCTGACTTA 243

QY 420 CTAAGGCTATTCAAGAACAAATTTGATCGTCCACAGTACAGACGACTACTTTGAG 479
DB 244 CTAAGGCTATTCAAGAACAAATTTGATCGTCCACAGTACAGACGACTACTTTGAG 303

QY 480 GTCAATGATTTTGAAGCGATGCAACCATTAAGTATGATGCAACCGCAGGTCTACTTTGCT 539
DB 304 GTCAATGATTTTGAAGCGATGCAACCATTAAGTATGATGCAACCGCAGGTCTACTTTGCT 363
```


Qy	540	GACAAAGATGGTTGGTAACTTGGCGAACCCACCTGTCCAGAAATTTTGGCTAAGCGGA	599
Db	364	GACAAAGATGGTTGGTAACTTGGCGAACCCACCTGTCCAGAAATTTTGGCTAAGCGGA	423
Qy	600	CATGTGCGGTTAGACCATATAAGAAAAACCATAACAAACCAAGCGAAATCTGTGAT	659
Db	424	CATGTGCGGTTAGACCATATAAGAAAAACCATAACAAACCAAGCGAAATCTGTGAT	483
Qy	660	GTGGGATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGACGAGTCTC	719
Db	484	GTGGGATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGACGAGTCTC	543
Qy	720	AAAGATATACGATTTGAAAAACATAGCTATCCGGTGACACCATCATCTCAAGAAATTA	779
Db	544	AAAGATATACGATTTGAAAAACATAGCTATCCGGTGACACCATCATCTCAAGAAATTA	603
Qy	780	CTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT	839
Db	604	CTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT	663
Qy	840	GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAA	899
Db	664	GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAA	723
Qy	900	GAGTTTACTTTACCGTGTAAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGT	959
Db	724	GAGTTTACTTTACCGTGTAAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGT	783
Qy	960	CTGAATGAAGAAATAAACAACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA	1019
Db	784	CTGAATGAAGAAATAAACAACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA	843
Qy	1020	GGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACTGTTTACCATCAAAATAC	1079
Db	844	GGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACTGTTTACCATCAAAATAC	903
Qy	1080	GTTGATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGACGT	1139
Db	904	GTTGATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGACGT	963
Qy	1140	AACTTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGCTTAAACTCTTACAACAAT	1199
Db	964	AACTTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGCTTAAACTCTTACAACAAT	1023
Qy	1200	CTCGATGCTTTTGGTATATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGAT	1259
Db	1024	CTCGATGCTTTTGGTATATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGAT	1083
Qy	1260	GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGGATGCTAGC	1319
Db	1084	GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGGATGCTAGC	1143
Qy	1320	TATCATTTAGCCGGTGGT 1337	
Db	1144	TATCATTTAGCCGTATGAT 1161	
RESULT 13			
LOCUS	E01413	1407 bp	DNA linear PAT 29-SEP-1997
DEFINITION	DNA sequence of streptokinase.		
ACCESSION	E01413		
VERSION	E01413.1	GI:2169669	
KEYWORDS	JP 1987296881-A/1.		
SOURCE	Lycopersicon esculentum (Solanum lycopersicum)		
ORGANISM	Lycopersicon esculentum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.		
REFERENCE	1 (bases 1 to 1407)		
AUTHORS	Marii, J.H. and Deebuitsudo, U.S.		
TITLE	YEAST DNA FRAGMENT AND ITS UTILIZATION		
JOURNAL	Patent: JP 1987296881-A 1 24-DEC-1987;		
COMMENT	PHILLIPS PETROLEUM CO		
PN	JP 1987296881-A/1		
PD	24-DEC-1987		
PF	02-MAY-1987 JP 1987109620		
PR	08-MAY-1986 US 86 860960		
PI	MARII JIEIN HEIJENSON, DEEBUITSUDO UOOMATSUKU SUTROOMAN PC		
	C12N15/00, C12N1/16, C12N9/12, (C12N1/16, C12R1/84); CC		
	strandedness:		
	Double;		
CC	topology: Linear;		
CC	hypothetical: No;		
CC	anti-sense: No;		
PH	Key		
FT	Location/Qualifiers		
FT	CDS		
FT	8..1255		
FT	/product='streptokinase'		
FT	1..7		
FT	5'UTR		
FT	1256..1407.		
FT	3'UTR		
FEATURES	Location/Qualifiers		
source	1..1407		
	/organism="Lycopersicon esculentum"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:4081"		
ORIGIN			
	Query Match 67.8%; Score 1126.2; DB 6; Length 1407;		
	Best Local Similarity 98.0%; Pred. No. 1.5e-285; Indels 0; Gaps 0;		
	Matches 1140; Conservative 0; Mismatches 23;		
Qy	175	ATACCATGATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCCAAT	234
Db	3	ATTCCATGATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCCAAT	62
Qy	235	TGGTTGTTAGGTTGCTGGTACTGTTGAGGGGACGAATCAAGAATTAGTCTTAAATTTT	294
Db	63	TAGTTGTTAGGTTGCTGGTACTGTTGAGGGGACGAATCAAGAATTAGTCTTAAATTTT	122
Qy	295	TTGAAATCGATCTAAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAATC	354
Db	123	TTGAAATCGATCTAAACATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAATC	182
Qy	355	CAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTG	414
Db	183	CAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTG	242
Qy	415	ACTTACTTAAAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAACGACACT	474
Db	243	ACTTACTTAAAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAACGACACT	302
Qy	475	TTGAGGTCATTTGTTTGGCAAGCGATGCAACCATTTACTGATCGGAAACGGCAAGGCT	534
Db	303	TTGAGGTCATTTGTTTGGCAAGCGATGCAACCATTTACTGATCGGAAACGGCAAGGCT	362
Qy	535	TTGCTGCAAAAGATGGTTCCGTAACTTGGCGACCCAACTGTCCAAAGAAATTTTGTGTA	594
Db	363	TTGCTGCAAAAGATGGTTCCGTAACTTGGCGACCCAACTGTCCAAAGAAATTTTGTGTA	422
Qy	595	GCAGCATGTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTG	654
Db	423	GCAGCATGTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTG	482
Qy	655	TTGATGCGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCC	714
Db	483	TTGATGCGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCC	542
Qy	715	GTCTCAAGATATCTAAGCTATTGAAAACACTTAGCTATCGGTGACACCATCACTCTCA	774
Db	543	GTCTCAAGATATCTAAGCTATTGAAAACACTTAGCTATCGGTGACACCATCACTCTCA	602
Qy	775	AATTACTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTTATACGATTTATG	834
Db	603	AATTACTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTTATACGATTTATG	662

```
QY 835 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGG 894
Db 663 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGG 722
QY 895 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAATAAT 954
Db 723 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAATAAT 782
QY 955 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATTACGTCCTTA 1014
Db 783 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATTACGTCCTTA 842
QY 1015 AAAAAGGGGAAAAGCCGATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCA 1074
Db 843 AAAAAGGGGAAAAGCCGATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCA 902
QY 1075 AATACGTTGATGTCGATACCAACAAATGCTTAAAGATGACGAGCTCTTAAACGATAGG 1134
Db 903 AATACGTTGATGTCGATACCAACAAATGCTTAAAGATGACGAGCTCTTAAACGATAGG 962
QY 1135 AACGTAACTTAGACTTTAGAGATTTATACGATCTCTGATGATAAGCTTAACTACTCTTACA 1194
Db 963 AACGTAACTTAGACTTTAGAGATTTATACGATCTCTGATGATAAGCTTAACTACTCTTACA 1022
QY 1195 ACAAATCTCGATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGGATATC 1254
Db 1023 ACAAATCTCGATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGGATATC 1082
QY 1255 ACAGTACACCAACCGTATACCAACCGTTTATATGGGCAAGCGACCGAAGGAGAGATG 1314
Db 1083 ACAGTACACCAACCGTATACCAACCGTTTATATGGGCAAGCGACCGAAGGAGAGATG 1142
QY 1315 CTAGCTATCATTTAGCGGTGGT 1337
Db 1143 CTAGCTATCATTTAGCGTATGAT 1165
```

```
RESULT 14
LOCUS A20016
DEFINITION A20016 1512 bp DNA linear PAT 14-JUL-1995
          alpha-factor.
ACCESSION A20016
VERSION A20016.1 GI:1247850
KEYWORDS synthetic construct
SOURCE synthetic construct
          other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1512)
          PROTEINS AND NUCLEIC ACIDS
          Patent: WO 9109125-A 24 27-JUN-1991;
FEATURES
          Location/Qualifiers
            1..1512
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
            7..1506
              /codon_start=1
              /transl_table=11
              /product="streptokinase fused to yeast alpha-factor"
              /protein_id="CAA01487.1"
              /db_xref="GI:1247851"
              /translation="MRPFSIFTVAVLFAASSALAAPVNTTETETAQIPABAVIGYLDL
EGDFVAVLFFSNTNGLLFINNTIASIAAKEEGVSLDKRIAGPEWLLDRPSVNSQ
LVSVAGTVEGTNODLSLAFPRIDLTSPAHGKTEQGLSPKSPATDSGAMPHKLE
KADLLKAOEQLIANVSHNDVFEVIDPASDAITDRNGKVYPADKDGSVLTPTQPVQ
EFLLSGHVRPPIKEPIQNOAKSVDEVTYVQTPINLPDDDFRPLKDTKLLKTLAIG
DTTIOELLAQASLKNKHPGTIYVERSSIVTHNDNIFRTILPMDQEFYHVKRE
QAYENKKSGLNEENINTDLISEKTVLKKGEKPPFPORSHLKLFTIKYVDVNTNEL
LKSEOLLTASERNLDPRLYDPRDKAKLLYNLDADFGLMDYTLTGKVEDNHDNTRI
TYVMGKRPGENASYHLAYDKORYTEEREVSYLRYTGTPIPDNPDK"
```

ORIGIN

```
Query Match
Best Local Similarity 67.8%; Score 1126.2; DB 6; Length 1512;
Matches 1140; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 175 ATACCATGATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAAT 234
Db 254 ATAAAGAAATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAAT 313
QY 235 TGGTTGTTAGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
Db 314 TAGTTGTTAGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 373
QY 295 TTGAATCGATCTAATCATCAACGCTGCTCATGAGGAAAGACAGACGAGCAAGCTTTAAGTC 354
Db 374 TTGAATTTGACCTTAACATCAACGCTGCTCATGAGGAAAGACAGACGAGCAAGCTTTAAGTC 433
QY 355 CAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAAGAAGCTG 414
Db 434 CAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAGAAGCTG 493
QY 415 ACTTACTAAAGGCTATTCAAGAACTAATTGATCGCTAAACGTCCACAGTAACGACGACTACT 474
Db 494 ACTTACTAAAGGCTATTCAAGAACTAATTGATCGCTAAACGTCCACAGTAACGACGACTACT 553
QY 475 TTGAGGTCATTTGATTTTGCAGACGATGCAACCATTTACTGATCGAAGCGCAAGCTTACT 534
Db 554 TTGAGGTCATTTGATTTTGCAGACGATGCAACCATTTACTGATCGAAGCGCAAGCTTACT 613
QY 535 TTGCTGACAAAGATGTTTCGGTAACTTTGCGGACCCCAACCTGTGCCAAGAAATTTTGTAA 594
Db 614 TTGCTGACAAAGATGTTTCGGTAACTTTGCGGACCCCAACCTGTGCCAAGAAATTTTGTAA 673
QY 595 GCGGACATGTCGGCTTAGACCATATAAAGAAAAAACAATCAAAACCAAGCGAAATCTG 654
Db 674 GCGGACATGTCGGCTTAGACCATATAAAGAAAAAACAATCAAAACCAAGCGAAATCTG 733
QY 655 TTGATGTGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAG 714
Db 734 TTGATGTGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAG 793
QY 715 GTCTCAAGATACTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAG 774
Db 794 GTCTCAAGATACTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAG 853
QY 775 AATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATG 834
Db 854 AATTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATG 913
QY 835 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGG 894
Db 914 AACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGG 973
QY 895 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAATAAT 954
Db 974 ATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGGATCAATAAATAAT 1033
QY 955 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATTACGTCCTTA 1014
Db 1034 CTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATTACGTCCTTA 1093
QY 1015 AAAAAGGGGAAAAGCCGATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCA 1074
Db 1094 AAAAAGGGGAAAAGCCGATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCA 1153
QY 1075 AATACGTTGATGTCGATACCAACAAATGCTTAAAGATGACGAGCTCTTAAACGATAGG 1134
Db 1154 AATACGTTGATGTCGATACCAACAAATGCTTAAAGATGACGAGCTCTTAAACGATAGG 1213
QY 1135 AACGTAACTTAGACTTTAGAGATTTATACGATCTCTGATGATAAGCTTAACTACTCTTACA 1194
Db 1214 AACGTAACTTAGACTTTAGAGATTTATACGATCTCTGATGATAAGCTTAACTACTCTTACA 1273
```

```

Qy 1195 ACAATCTCGATGCTTTGGTATTATGAGCTATACCTTTAACTGGAAAGTAGAGGATAATC 1254
Db 1274 ACAATCTCGATGCTTTGGTATTATGAGCTATACCTTTAACTGGAAAGTAGAGGATAATC 1333
Qy 1255 ACGATGACACCAACCGTATCATACCAACGTTTATATGGCGAAGCGACCCGAAGGAGGAATG 1314
Db 1334 ACGATGACACCAACCGTATCATACCAACGTTTATATGGCGAAGCGACCCGAAGGAGGAATG 1393
Qy 1315 CTAGCTATCATTTAGCCGGTGGT 1337
Db 1394 CTAGCTATCATTTAGCCCTATGAT 1416

RESULT 15
LOCUS I13204 1512 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 27 from patent US 5434073.
ACCESSION I13204
VERSION I13204.1 GI:910552
KEYWORDS
SOURCE
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1512)
AUTHORS Dawson,K., Hunter,M.G. and Czaplowski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 27 18-JUL-1995;
FEATURES
source Location/Qualifiers
1..1512
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 67.8%; Score 1126.2; DB 6; Length 1512;
Best Local Similarity 98.0%; Pred. No. 1.5e-285;
Matches 1140; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 175 ATACCATGATTCGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAT 234
Db 254 ATAAAGAAATTCGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAT 313
Qy 235 TCGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 294
Db 314 TAGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 373
Qy 295 TTGAATCGATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACCAAGGCTTAAGTC 354
Db 374 TTGAATTCGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGACCAAGGCTTAAGTC 433
Qy 355 CAAATCAAAACCCNTTGGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTG 414
Db 434 CAAATCAAAACCCNTTGGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTG 493
Qy 415 ACTTACTAAAGGCTATTCAAGAACCAATTCATCCCTAACGTCACAGTAACGACGACTACT 474
Db 494 ACTTACTAAAGGCTATTCAAGAACCAATTCATCCCTAACGTCACAGTAACGACGACTACT 553
Qy 475 TTGAGTCAATGATTTTCAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTTACT 534
Db 554 TTGAGTCAATGATTTTCAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTTACT 613
Qy 535 TTGCTGACAAAGATGGTTGGTAACTTGGCGACCCCACTGTCCAAAGAAATTTTGTCTAA 594
Db 614 TTGCTGACAAAGATGGTTGGTAACTTGGCGACCCCACTGTCCAAAGAAATTTTGTCTAA 673
Qy 595 GCGGACATGTCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTG 654
Db 674 GCGGACATGTCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTG 733
Qy 655 TTGATGTGGAATATACTGTACAGTTTACTCCCTTTAAACCCCTGTATGACGATTTTCAGACCAG 714
Db 734 TTGATGTGGAATATACTGTACAGTTTACTCCCTTTAAACCCCTGTATGACGATTTTCAGACCAG 793

```

```

Qy 715 GTCTCAAAGTACTTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTCATCTCAAG 774
Db 794 GTCTCAAAGTACTTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTCATCTCAAG 853
Qy 775 AATTACTAGCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATG 834
Db 854 AATTACTAGCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATG 913
Qy 835 AACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGG 894
Db 914 AACGTGACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGG 973
Qy 895 ATCAAGAGTTTACTTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAT 954
Db 974 ATCAAGAGTTTACTTTACCATGTCAAAATCGGGAACAAAGCTTTATGAGATCAATAAAAAT 1033
Qy 955 CTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 1014
Db 1034 CTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTA 1093
Qy 1015 AAAAAGGGGAAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACATCA 1074
Db 1094 AAAAAGGGGAAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACATCA 1153
Qy 1075 AATACGTTGATGTCGATACCAACCAAGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGTAGCG 1134
Db 1154 AATACGTTGATGTCGATACCAACCAAGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGTAGCG 1213
Qy 1135 AACGTAACTTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTTAACTACTCTACA 1194
Db 1214 AACGTAACTTAGACTTCAGAGATTTATACGATCTCTCGTGATAAGGCTTAACTACTCTACA 1273
Qy 1195 ACAATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAAGTAGAGGATAATC 1254
Db 1274 ACAATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAAGTAGAGGATAATC 1333
Qy 1255 ACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGGAATG 1314
Db 1334 ACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGGAATG 1393
Qy 1315 CTAGCTATCATTTAGCCGGTGGT 1337
Db 1394 CTAGCTATCATTTAGCCCTATGAT 1416

```

Search completed: February 1, 2006, 04:17:47
Job time : 8445.12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 1004.11 Seconds
(without alignments)
11024.772 Million cell updates/sec

Title: US-09-940-235-10

Perfect score: 1661

Sequence: 1 gcaaccccgccagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	100.0	1661	3	Aaa37637 Chimeric
2	1282.2	77.2	1377	3	Aaa37622 Streptoki
3	1219.4	73.4	1327	3	Aaa37628 Streptoki
4	1185	71.3	1541	3	Aaa37644 Chimeric
5	1165.8	70.2	2096	3	Aaa37643 Chimeric
6	1150.2	69.2	1782	3	Aaa37642 Chimeric
7	1150	69.2	2385	2	Aax80497 Streptoki
8	1148.8	69.2	2030	2	Aaa11651 FB-PB-SK
9	1148.6	69.2	1242	2	Aax80492 Streptoco
10	1148.6	69.2	1245	3	Aaa37633 S. equisi
11	1148.6	69.2	1254	6	ABA05546 Streptoki
12	1148.6	69.2	8893	6	ABA05547 Maxadilan
13	1147	69.1	1242	2	Aax16632 Streptoco
14	1143.8	68.9	1242	2	Aax16633 Streptoco
15	1136.2	68.4	7057	12	ADM01294 Plasmid p
16	1134.6	68.3	2566	2	Aat77778 Coding se
17	1126.2	67.8	1407	1	Aan70106 DNA encod
18	1126.2	67.8	1512	2	Aaa12158 Streptoki
19	1124.6	67.7	1245	2	Aaa20665 SKC-2 str

20	1124.6	67.7	1335	2	AAQ12156
21	1124.6	67.7	1458	2	AAQ12162
22	1124.6	67.7	1467	2	AAQ12490
23	1124.6	67.7	2589	2	AAQ12160
24	1116.6	67.2	1473	2	AAQ05603
25	1108.6	66.7	2568	1	AAN50493
26	1107	66.6	1323	2	AAAT29961
27	1099.4	66.2	1242	5	AAF82144
28	1096.6	66.0	1122	3	AAZ99251
29	1096.6	66.0	1158	3	AAZ99252
30	1089.2	65.6	1209	3	AAZ99249
31	1089.2	65.6	1245	3	AAZ99250
32	1080.4	65.0	2253	2	AAQ12161
33	1078.8	64.9	1119	2	AAQ12159
34	995	59.9	1320	6	ABN70192
35	995	59.9	1323	13	ADR83811
36	990.2	59.6	1473	2	AAQ05604
37	976.2	58.8	1068	2	AAx80493
38	971.6	58.5	2208	2	AAx83589
39	899.4	54.1	1245	10	ADF48644
40	762.8	45.9	1262	2	AAQ10230
41	450	27.1	450	2	AAx80494
42	316.8	19.1	4860	3	AAA35009
43	316.8	19.1	6988	9	ACD06169
44	316.8	19.1	7020	13	ADR90518
45	316.8	19.1	7049	12	ADN95947

ALIGNMENTS

RESULT 1

AAA37637
ID AAA37637 standard; DNA; 1661 BP.

XX AC AAA37637;

XX DT 15-SEP-2003 (revised)

XX DT 13-OCT-2000 (first entry)

XX DE Chimeric SK-FBD coding sequence.

XX KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
XX KW plasminogen; human; fibronectin; thrombolytic therapy;
XX KW cardiovascular disorder; ss.

XX OS Streptococcus dysgalactiae subsp. equisimilis.
XX OS Homo sapiens.
XX OS Chimeric.

XX PN EPI024192-A2.

XX PD 02-AUG-2000.

XX PF 23-DEC-1999; 99EP-00310541.

XX PR 24-DEC-1998; 98IN-DE003825.

XX PA (COUL) CSIR COUNCIL SCI IND RES.

XX PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
XX PI Yadav M;

XX DR WPI; 2000-516032/47.

XX PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
XX PT domains of human fibronectin.

XX EX Example 5; Fig 19b; 58pp; English.

XX CC This sequence represents a chimeric streptokinase-fibrin binding domain
XX CC (SK-FBD) protein coding sequence. The invention relates to a hybrid

DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX Streptokinase-NTRN gene.
 DE
 XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 OS
 XX EP1024192-A2.
 PN
 XX 02-AUG-2000.
 PD
 XX 23-DEC-1999; 99EP-00310541.
 XX
 XX 24-DEC-1998; 98IN-DE003825.
 PR
 XX (COUL) CSIR COUNCIL SCI IND RES.
 PA
 XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 PI
 XX WPI; 2000-516032/47.
 DR
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 PS Example 1; Fig 11; 58pp; English.
 XX
 CC This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN
 CC stands for N-terminally repaired with native sequence). The invention
 CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
 CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
 CC activation, and fibrin binding regions of human fibrinectin, which are
 CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
 CC possesses the ability to bind with fibrin independently and also
 CC characteristically retains a PG activation ability which becomes evident
 CC only after a pronounced duration, or lag, after exposure of the PA to a
 CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
 CC domain polypeptides are useful in thrombolytic therapy for various kinds
 CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
 CC as well as kinetics of plasminogen activation that are distinct from that
 CC of natural streptokinase in being characterised by a temporary delay, or
 CC lag of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;
 Query Match 77.2%; Score 1282.2; DB 3; Length 1377;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1284; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 51 GCACCCGTCGCGAGACCAACGCTCCGAGATCTCGATCCGCGGAAATTAATACGACT 110
 Db 1 GCACCCGTCGCGAGACCAACGCTCCGAGATCTCGATCCGCGGAAATTAATACGACT 60
 Qy 111 CACTATAGGAGACCAACGCTCCCTCTAGAAATAATTTGTTTAACTTTAAGAGG 170
 Db 61 CACTATAGGAGACCAACGCTCCCTCTAGAAATAATTTGTTTAACTTTAAGAGG 120
 Qy 171 AGATATACCATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAACAGC 230
 Db 121 AGATATACCATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACCAACAGC 180

Qy 231 CAATTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 290
 Db 181 CAATTGGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAA 240
 Qy 291 TTTTGTGAATTCGATCTAAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTA 350
 Db 241 TTTTGTGAATTCGATCTAAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTA 300
 Qy 351 AGTCCAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGTCATCATATAAATCTTCAGAAA 410
 Db 301 AGTCCAAAATCAAAAACCAATTTGCTACTGATAGTGGCGGATGTCATCATATAAATCTTCAGAAA 360
 Qy 411 GCTGACTTTACTAAAGGCTTATTCAAGAACCAATTTGATCGCTAACGTCACAGTAACCAACGAC 470
 Db 361 GCTGACTTTACTAAAGGCTTATTCAAGAACCAATTTGATCGCTAACGTCACAGTAACCAACGAC 420
 Qy 471 TACTTTGAGGTCATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTC 530
 Db 421 TACTTTGAGGTCATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTC 480
 Qy 531 TACTTTGCTGCAAAAGATGTTTCGGTAACTTGGCGCCACCTGTCGAAAGATTTTTCG 590
 Db 481 TACTTTGCTGCAAAAGATGTTTCGGTAACTTGGCGCCACCTGTCGAAAGATTTTTCG 540
 Qy 591 CTAAGCGGACATGTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAA 650
 Db 541 CTAAGCGGACATGTGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAA 600
 Qy 651 TCTGTTGATGTGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGA 710
 Db 601 TCTGTTGATGTGGAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGA 660
 Qy 711 CCAGGTCCTCAAGATACCTAACTATTGAAAAACATAGCTATCGGTGACACCATCATCT 770
 Db 661 CCAGGTCCTCAAGATACCTAACTATTGAAAAACATAGCTATCGGTGACACCATCATCT 720
 Qy 771 CAAGAAATTACTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCAACAGGCTTATACGATT 830
 Db 721 CAAGAAATTACTAGCTCAAGCAAAAGCAATTTTAAACAAAAACCAACAGGCTTATACGATT 780
 Qy 831 TATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCA 890
 Db 781 TATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCA 840
 Qy 891 ATGATCAAGATTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAA 950
 Db 841 ATGATCAAGATTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAA 900
 Qy 951 AAATCTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTC 1010
 Db 901 AAATCTGCTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTC 960
 Qy 1011 CTTAAAAAAGGGGAAAGCCGTATGATCCCTTTCATCGCAGTCACTTCAAACTGTTTACC 1070
 Db 961 CTTAAAAAAGGGGAAAGCCGTATGATCCCTTTCATCGCAGTCACTTCAAACTGTTTACC 1020
 Qy 1071 ATCAAAATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCT 1130
 Db 1021 ATCAAAATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCT 1080
 Qy 1131 AGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCTCTGATTAAGCTTAAACTACTC 1190
 Db 1081 AGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCTCTGATTAAGCTTAAACTACTC 1140
 Qy 1191 TACAACCAATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACTTGGAAAGTAGAGGAT 1250
 Db 1141 TACAACCAATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACTTGGAAAGTAGAGGAT 1200
 Qy 1251 AATCAGCATGACCAACCGCTATCATAAACCGTTTATATGGCAACGACCGCAAGGAGAG 1310
 Db 1201 AATCAGCATGACCAACCGCTATCATAAACCGTTTATATGGCAACGACCGCAAGGAGAG 1260
 Qy 1311 AATGCTAGCTATCATTTTAGCCGGTGGT 1337

QY 1241 ACTAGAGGATATCAGATGACACCAACCGTATCATAAACCGTTATATGGCGAGCGACC 1300
 Db |||||
 QY 1301 CGAAGGAGAGATGCTAGCTATCATTTAGCGGTGGT 1337
 Db |||||

RESULT 4
 AAA37644
 ID AAA37644 standard; DNA; 1541 BP.
 XX
 AC AAA37644;
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE Chimeric SK-FBD coding sequence.
 XX
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Homo sapiens.
 OS Chimeric.
 XX
 PN EP1024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RES.
 XX
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX
 DR WPI; 2000-516032/47.
 XX
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 PS Disclosure; Fig 17b; 58pp; English.
 XX
 CC This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

SQ Sequence 1541 BP; 497 A; 328 C; 335 G; 381 T; 0 U; 0 Other;
 Query Match 71.3%; Score 1185; DB 3; Length 1541;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 151 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 210
 Db 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATAGTGGTCTCTGAATGGCTACTAGATC 60
 QY 211 GTCCATCTGTCAACACAGCCAAATTTGGTTGTAGCGTTGCTGGTACTGTGTAGGGAGCA 270
 Db |||||
 QY 61 GTCTTCTGTAAATAACACGCCAAATTTGGTTGTAGCGTTGCTGGTACTGTGTAGGGAGCA 120
 Db |||||
 QY 271 ATCAAGCATTAGTCTTAAATTTTTGAAATCGATCTAACATCAGCATCTCATGGAG 330
 Db |||||
 QY 121 ATCAAGCATTAGTCTTAAATTTTTGAAATCGATCTAACATCAGCATCTCATGGAG 180
 Db |||||
 QY 331 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGCA 390
 Db |||||
 QY 181 GAAAGACAGAGCAAGGCTTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGCA 240
 Db |||||
 QY 391 TGTCACTAAACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAAATGTATGCTGA 450
 Db |||||
 QY 241 TGTCACTAAACTTTGAGAAAGCTGACTTAAAGGCTATTCAAGAAACAAATGTATGCTGA 300
 Db |||||
 QY 451 ACGTCCACAGTAACGAGCATCTTTGAGGTCATTGATTTTGCACGCGATGCAACCATTA 510
 Db |||||
 QY 301 ACGTCCACAGTAACGAGCATCTTTGAGGTCATTGATTTTGCACGCGATGCAACCATTA 360
 Db |||||
 QY 511 CTGATCGAAACCGCAAGGCTTACTTTGCTGCAAAAGATGGTTTCGGTAACTTTGCCGACCC 570
 Db |||||
 QY 361 CTGATCGAAACCGCAAGGCTTACTTTGCTGCAAAAGATGGTTTCGGTAACTTTGCCGACCC 420
 Db |||||
 QY 571 AACCTGTCCAAGAAATTTTGTCAAGCGGACATGTGCGGTTTAGACCATATATAAGAAAC 630
 Db |||||
 QY 421 AACCTGTCCAAGAAATTTTGTCAAGCGGACATGTGCGGTTTAGACCATATATAAGAAAC 480
 Db |||||
 QY 631 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAAATATCTGTACAGTTTACTCCTTTAA 690
 Db |||||
 QY 481 CAATACAAAACCAAGCGAAATCTGTTGATGTGGAAATATCTGTACAGTTTACTCCTTTAA 540
 Db |||||
 QY 691 ACCCTGATGACGATTTTCAGACCGAGTCTCAAAGATACTTAAGTATTTGAAACACATAGCTA 750
 Db |||||
 QY 541 ACCCTGATGACGATTTTCAGACCGAGTCTCAAAGATACTTAAGTATTTGAAACACATAGCTA 600
 Db |||||
 QY 751 TCGGTGACACCATCACATCTCAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACAAA 810
 Db |||||
 QY 601 TCGGTGACACCATCACATCTCAAGAAATTTACTAGCTCAAGCAACAAGCAATTTTAAACAAA 660
 Db |||||
 QY 811 ACCACCGAGCTATACGATTTTGAACGAGTCTCCTCAATCGTCACTCATGACCAATGACA 870
 Db |||||
 QY 661 ACCACCGAGCTATACGATTTTGAACGAGTCTCCTCAATCGTCACTCATGACCAATGACA 720
 Db |||||
 QY 871 TTTTCCGTGAGTTTTCACCAATGATCAAGAGTTTACTTACCGTGTTTAAAAATCGGNAAC 930
 Db |||||
 QY 721 TTTTCCGTGAGTTTTCACCAATGATCAAGAGTTTACTTACCGTGTTTAAAAATCGGNAAC 780
 Db |||||
 QY 931 AAGCTTATAGCATCAATATAAAATTTCTGCTCTGAATGAAGAAATATAACCAACGATGACA 990
 Db |||||
 QY 781 AAGCTTATAGCATCAATATAAAATTTCTGCTCTGAATGAAGAAATATAACCAACGATGACA 840
 Db |||||
 QY 991 TCTCTGAGAAATATTACGCTCTTAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCA 1050
 Db |||||
 QY 841 TCTCTGAGAAATATTACGCTCTTAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCA 900
 Db |||||
 QY 1051 GTCACTTGAACCTGTTCCACCATCAATACGTTGATGTCGATACCAACGAATTTGCTAAAAA 1110
 Db |||||
 QY 901 GTCACTTGAACCTGTTCCACCATCAATACGTTGATGTCGATACCAACGAATTTGCTAAAAA 960
 Db |||||
 QY 1111 GTGAGCAGCTCTTAAACAGCTAGCGAAAGTAACTTAGACTTTTACAGAGATTTTATACCATCCTC 1170
 Db |||||
 QY 961 GTGAGCAGCTCTTAAACAGCTAGCGAAAGTAACTTAGACTTTTACAGAGATTTTATACCATCCTC 1020
 Db |||||

QY 1171 GTGATAAGGCTAACTACTCTACAACTCTCGATGCTTTTGGTATTATGGACTATACCT 1230
 Db |||||
 QY 1021 GTGATAAGGCTAACTACTCTACAACTCTCGATGCTTTTGGTATTATGGACTATACCT 1080
 Db |||||
 QY 1231 TAACTGGAAGGTAGAGGATAATACGATGACACCAACCGTATCATACCGTTTATATGG 1290
 Db |||||
 QY 1081 TAACTGGAAGGTAGAGGATAATACGATGACACCAACCGTATCATACCGTTTATATGG 1140
 QY 1291 GCAAGCGACCCGAGAGAGATGCTAGCTATCATTTAGCGGTGGTGTGACGGCGAGC 1350
 Db |||||
 QY 1141 GCAAGCGACCCGAGAGAGATGCTAGCTATCATTTAGCTGCTGGTGGCGCGCAAC 1200
 QY 1351 AAATGGTTC 1359
 Db |||||
 QY 1201 AGATTGTAC 1209

RESULT 5
 AAA37643
 ID AAA37643 standard; DNA; 2096 BP.
 XX

AC AAA37643;

DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)

XX Chimeric SK-FBD coding sequence.

DE Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibronectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS Homo sapiens.

OS Chimeric.

XX EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99BP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;

XX WPI; 2000-516032/47.
 DR
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibronectin.

XX Example 6; Fig 22b; 58pp; English.

XX This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (Pg) activation, and
 CC fibrin binding regions of human fibronectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of

CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulation of
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX
 SQ Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;

Query Match 70.2%; Score 1165.8; DB 3; Length 2096;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCGCAATTTGGTTGTT 242
 Db |||||
 QY 588 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCGCAATTTGGTTGTT 647
 QY 243 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATTTTGAATTC 302
 Db |||||
 QY 648 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATTTTGAATTC 707
 QY 303 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 362
 Db |||||
 QY 708 GATCTAAATCAACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 767
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAATTGAGAAAGCTGACTTACTA 422
 Db |||||
 QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAATTGAGAAAGCTGACTTACTA 827
 QY 423 AAGGCTATTCAAGAAACAATTGATCGCTAAACCTCCACAGTACACGACGACTCTTTGAGTGC 482
 Db |||||
 QY 828 AAGGCTATTCAAGAAACAATTGATCGCTAAACCTCCACAGTACACGACGACTCTTTGAGTGC 887
 QY 483 ATTGATTTTGCAGCGGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 542
 Db |||||
 QY 888 ATTGATTTTGCAGCGGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 947
 QY 543 AAAGATGTTTGGTAAACCTTGGCCGACCACTCTGTCAGAGAAATTTTGGCTAAGCGGACAT 602
 Db |||||
 QY 948 AAAGATGTTTGGTAAACCTTGGCCGACCACTCTGTCAGAGAAATTTTGGCTAAGCGGACAT 1007
 QY 603 GTGCGCGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db |||||
 QY 1008 GTGCGCGTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1067
 QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGACAGGCTCTCAA 722
 Db |||||
 QY 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGACAGGCTCTCAA 1127
 QY 723 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATTCACATCTCAAGAAATTA 782
 Db |||||
 QY 1128 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATTCACATCTCAAGAAATTA 1187
 QY 783 GCTCAAGCACAAGCATTTTAAACCAACCAACCAAGCGCTATAGTATTATGAAAGCTGAC 842
 Db |||||
 QY 1188 GCTCAAGCACAAGCATTTTAAACCAACCAACCAAGCGCTATAGTATTATGAAAGCTGAC 1247
 QY 843 TCCTCAATCTGCTACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db |||||
 QY 1248 TCCTCAATCTGCTACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1307
 QY 903 TTTACTTACCGTGTATAAATTCGGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db |||||
 QY 1308 TTTACTTACCGTGTATAAATTCGGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1367
 QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG 1022
 Db |||||
 QY 1368 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG 1427
 QY 1023 GAAAGCGGTATGATCCCTTTTGTGCGAGTGCACCTTGTAAACCTGTTTCAACCATCAATACGTT 1082
 Db |||||
 QY 1428 GAAAGCGGTATGATCCCTTTTGTGCGAGTGCACCTTGTAAACCTGTTTCAACCATCAATACGTT 1487

QY 1083 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
D 1488 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1547
QY 1143 TTAGACTTCAGAGATTTATAGACTTCCTCGTGATAGGCTAACTACTCTACACAAATCTC 1202
D 1548 TTAGACTTCAGAGATTTATAGACTTCCTCGTGATAGGCTAACTACTCTACACAAATCTC 1607
QY 1203 GATGCTTTTGGTATTATGGACTACTACTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
D 1608 GATGCTTTTGGTATTATGGACTACTACTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1667
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1322
D 1668 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1727
QY 1323 CATTTAGCGGTGGTGGTTCAGGGCGCAGCAAAATGGTTTC 1359
D 1728 CATTTAGCTGGTGGTGGCCAGCGCAACAGATTGTAC 1764

RESULT 6

AAA37642
ID AAA37642 standard; DNA; 1782 BP.
XX
AC AAA37642;
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
Chimeric SK-FBD coding sequence.
XX
Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
EP1024192-A2.
XX
02-AUG-2000.
XX
23-DEC-1999; 99EP-00310541.
XX
24-DEC-1998; 98IN-DE003825.
XX
(COUL) CSIR COUNCIL SCI IND RES.
XX
Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
WPI; 2000-516032/47.
XX

PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibronectin.
XX
XX Example 5; Fig 21b; 58pp; English.
XX

CC This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibronectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of

CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Query Match 69.2%; Score 1150.2; DB 3; Length 1782;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTTGT 242
D 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTTGT 597
QY 243 AGCGTTGCTGTACTGTTGAGGGAGCAATCAAGACATTAGTCTTTAAATTTTGAATC 302
D 598 AGCGTTGCTGTACTGTTGAGGGAGCAATCAAGACATTAGTCTTTAAATTTTGAATC 657
QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGTTAAGTCCAAATCA 362
D 658 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGTTAAGTCCAAATCA 717
QY 363 AAACCATTTGCTACTGTAGTGGCGGATGTACATAAATTTGAGAAAGCTGACTTACTA 422
D 718 AAACCATTTGCTACTGTAGTGGCGGATGTACATAAATTTGAGAAAGCTGACTTACTA 777
QY 423 AAGGCTATTCAAGAACCAATTTGATCGTTAAACGTCCACAGTACGACGACTACTTTGAGTGC 482
D 778 AAGGCTATTCAAGAACCAATTTGATCGTTAAACGTCCACAGTACGACGACTACTTTGAGTGC 837
QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTTCTATTGCTGAC 542
D 838 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTTCTATTGCTGAC 897
QY 543 AAGCATGTTTCGGTAACTTCCCGACCCCACTGTCCAGAAATTTTTCGTAAGCGGACAT 602
D 898 AAGCATGTTTCGGTAACTTCCCGACCCCACTGTCCAGAAATTTTTCGTAAGCGGACAT 957
QY 603 GTGCGCGTTAGACCATATAAGAAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 662
D 958 GTGCGCGTTAGACCATATAAGAAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 1017
QY 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTTTCAGACCGGTTCTCAA 722
D 1018 GAATATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTTTCAGACCGGTTCTCAA 1077
QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGATTTACTA 782
D 1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGATTTACTA 1137
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGCGCTATAGATTTTATGAAAGCTGAC 842
D 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGCGCTATAGATTTTATGAAAGCTGAC 1197
QY 843 TCCTCAATCTGCTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
D 1198 TCCTCAATCTGCTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
QY 903 TTTACTTACCGTGTAAAAAATCGGGAACAAGTTATAGGATCAATAAAAAATCTGGTCTG 962
D 1258 TTTACTTACCGTGTAAAAAATCGGGAACAAGTTATAGGATCAATAAAAAATCTGGTCTG 1317
QY 963 AATGAGAAATAAACACACATGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAGGG 1022
D 1318 AATGAGAAATAAACACACATGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAGGG 1377

QY 1023 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCCAATCAATAGCTT 1082
 DB 1378 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCCAATCAATAGCTT 1437
 QY 1083 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 DB 1438 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
 QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGCTAAACTACTCTCAACAATCTC 1202
 DB 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGCTAAACTACTCTCAACAATCTC 1557
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGATATACAGATGAC 1262
 DB 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGATATACAGATGAC 1617
 QY 1263 ACCAACCGTATACATACCGTTTATATGGGCAACGCGACCGAAGGAGAGATGCTAGCTAT 1322
 DB 1618 ACCAACCGTATACATACCGTTTATATGGGCAACGCGACCGAAGGAGAGATGCTAGCTAT 1677
 QY 1323 CATTTAGCCGGTGGT 1337
 DB 1678 CATTTAGCCGTATGAT 1692

RESULT 7

AAx80497
 ID AAx80497 standard; cDNA; 2385 BP.
 XX
 AC AAx80497;
 XX
 DT 26-AUG-1999 (first entry)
 XX
 DE Streptokinase and maltose binding protein fusion protein encoding cDNA.
 XX
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW NSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 OS
 PN WO9931247-A1.
 XX
 PD 24-JUN-1999.
 XX
 PF 15-DEC-1998; 98WO-US026694.
 XX
 PR 15-DEC-1997; 97US-0069497P.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Reed GL;
 XX
 DR WPI; 1999-395183/33.
 DR P-PSDB; AAY24797.
 XX
 XX N-terminally deleted streptokinase.
 XX
 PS Example; Page 45-48; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein;
 CC nucleic acid (I) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a

CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC encodes a streptokinase and maltose binding protein fusion protein from
 CC an example of the present invention
 XX
 SQ Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;
 Query Match 69.2%; Score 1150; DB 2; Length 2385;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1153; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 180 ATGATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGGTT 239
 DB 1141 AGGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGGTT 1200
 QY 240 GTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAA 299
 DB 1201 GTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAA 1260
 QY 300 ATCGATCTAACATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAA 359
 DB 1261 ATCGATCTAACATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAA 1320
 QY 360 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCAATAAATTTGAGAAAGCTGACTTA 419
 DB 1321 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTCAATAAATTTGAGAAAGCTGACTTA 1380
 QY 420 CTTAAAGCTATTCAAGAAACAATTGATCGCTTAAACCTCCACAGTAACGACGACTACTTTGAG 479
 DB 1381 CTTAAAGCTATTCAAGAAACAATTGATCGCTTAAACCTCCACAGTAACGACGACTACTTTGAG 1440
 QY 480 GTCAATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGCT 539
 DB 1441 GTCAATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGCT 1500
 QY 540 GACAAAGATGCTTCCGTTAACTTCCGCGACCCCACTGTCCAAAGATTTTGTCTAAGCGGA 599
 DB 1501 GACAAAGATGCTTCCGTTAACTTCCGCGACCCCACTGTCCAAAGATTTTGTCTAAGCGGA 1560
 QY 600 CATGTGCGCTTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGAT 659
 DB 1561 CATGTGCGCTTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGAT 1620
 QY 660 GTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCCAGACGAGTCTC 719
 DB 1621 GTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCCAGACGAGTCTC 1680
 QY 720 AAGATATATCTGTGAAAAACCTAGCTATCGGTGACACCATCATCTCAAGAAATTA 779
 DB 1681 AAGATATATCTGTGAAAAACCTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1740
 QY 780 CTAGCTCAAGCACAAGACATTTTAAACAAAAACCAAGCGCTATACGATTTATGAAACGT 839
 DB 1741 CTAGCTCAAGCACAAGACATTTTAAACAAAAACCAAGCGCTATACGATTTATGAAACGT 1800
 QY 840 GACTCTCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAA 899
 DB 1801 GACTCTCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAA 1860
 QY 900 GAGTTTACTTACCGTGTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGT 959
 DB 1861 GAGTTTACTTACCGTGTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGT 1920
 QY 960 CTGATCAAGCAATAAACAACACCTGATCTCTGAGAAATATTAGTCTCTTAAAAAA 1019

Db 1921 CTGAATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAA 1980
Qy 1020 GGGGAAAGCCGATGATCCCTTTGATCGAGTCACCTTTGAAACTGTTCAACCATCAAAATAC 1079
Db 1981 GGGGAAAGCCGATGATCCCTTTGATCGAGTCACCTTTGAAACTGTTCAACCATCAAAATAC 2040
Qy 1080 GTTGATGTCATACCAACGAATTGCTTAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGT 1139
Db 2041 GTTGATGTCATACCAACGAATTGCTTAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGT 2100
Qy 1140 AACTTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTAACTACTCTACAACAT 1199
Db 2101 AACTTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTAACTACTCTACAACAT 2160
Qy 1200 CTGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATAATCAGAT 1259
Db 2161 CTGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATAATCAGAT 2220
Qy 1260 GACACCAACCGTATCATACCGTTTATGCGGCAAGCGACCCGAAAGGAGAGAAATGCTAGC 1319
Db 2221 GACACCAACCGTATCATACCGTTTATGCGGCAAGCGACCCGAAAGGAGAGAAATGCTAGC 2280
Qy 1320 TATCATTTAGCCGGTGGT 1337
Db 2281 TATCATTTAGCCCTATGAT 2298

RESULT 8
AAQ11651
ID AAQ11651 standard; DNA; 2030 BP.
XX
AC AAQ11651;
XX
DT 08-JUL-1991 (first entry)
XX
DE FB-FB-SK fusion construct.
XX
KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW streptokinase; fusion protein; ss.
XX
OS Staphylococcus aureus.
XX
FH Key
CDS 10..184
FT /tag= a
FT /label= FB monomer
FT 185..358
FT /tag= b
FT /label= FB monomer
FT 359..1601
FT /tag= c
FT /label= streptokinase
XX
PN US5011686-A.
XX
PD 30-APR-1991.
XX
PF 15-NOV-1989; 89US-00437769.
XX
PR 21-SEP-1987; 87US-00099242.
XX
PA (CREA-) CREATIVE BIOMOLEC.
XX
PI Pang RHL;
XX
DR WPI; 1991-140198/19.
DR P-PSDB; AAR11829.
XX
XX Imparting injectable fibrinolytic agent - with affinity for intravascular
XX thrombus, by linking agent to fibrin binding domain.
XX
PS Disclosure; Fig 5; 18pp; English.

CC The DNA encodes an FB-FB dimer linked to the streptokinase coding
CC sequence. The FB fragment has selective affinity for fibrin, low affinity
CC for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
CC capability. See also AAQ11649 and AAQ11650
XX
SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;

Query Match 69.2%; Score 1148.8; DB 2; Length 2030;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 178 COATGATTCGTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTGG 237
Db 353 CGAAGATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTAG 412
Qy 238 TTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTTTG 297
Db 413 TTGTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTTTG 472
Qy 298 AAATCGATCTAAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAA 357
Db 473 AAATCGATCTAAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTTAAGTCCAA 532
Qy 358 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACT 417
Db 533 AATCAAAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACT 592
Qy 418 TACTAAAGGCTATTCAAGAACAAATTTGATCGTCCACAGTCCACAGTAAAGCACTACTTTG 477
Db 593 TACTAAAGGCTATTCAAGAACAAATTTGATCGTCCACAGTCCACAGTAAAGCACTACTTTG 652
Qy 478 AGGTCATTGATTTTGCAGGCGATGCAACCATTTACTGATCGGAAACGGCAGGCTACTTTG 537
Db 653 AGGTCATTGATTTTGCAGGCGATGCAACCATTTACTGATCGGAAACGGCAGGCTACTTTG 712
Qy 538 CTGACAAAGATGTTTGGTAACTTGGCCGACCAACCTGTGCCAAGAAATTTTGTCTAAGCG 597
Db 713 CTGACAAAGATGTTTGGTAACTTGGCCGACCAACCTGTGCCAAGAAATTTTGTCTAAGCG 772
Qy 598 GACATGTCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTG 657
Db 773 GACATGTCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTG 832
Qy 658 ATGTGGAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTAGACAGGTC 717
Db 833 ATGTGGAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTAGACAGGTC 892
Qy 718 TCAAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAAT 777
Db 893 TCAAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAAT 952
Qy 778 TACTAGCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAC 837
Db 953 TACTAGCTCAAGCACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAC 1012
Qy 838 GTGACTCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATC 897
Db 1013 GTGACTCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATC 1072
Qy 898 AAGAGTTTACTTACCGGTGTTAAATAATCGGGAACAGCTTATAGGATCAATAAAAACTCG 957
Db 1073 AAGAGTTTACTTACCGGTGTTAAATAATCGGGAACAGCTTATAGGATCAATAAAAACTCG 1132
Qy 958 GTCTGAATGAAGAAATAAAACACTGACCTGTACTCTGAGAAATATTACGCTCTTAAAA 1017
Db 1133 GTCTGAATGAAGAAATAAAACACTGACCTGTACTCTGAGAAATATTACGCTCTTAAAA 1192
Qy 1018 AAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTGTCAACATCAAT 1077
Db 1193 AAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTGTCAACATCAAT 1252
Qy 1078 ACGTTGATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGGAC 1137

Db 1253 ACGTTGATCGATACCAACGAATTGCTGTAATAAGTGAGCAGCTCTTTAAACAGCTAGCGAAC 1312
 Qy 1138 GTAACCTAGACTTCAGAGATTTATAGATCCTCGTGATAGGCTAAACTACTCTACAAACA 1197
 Db 1313 GTAACCTAGACTTCAGAGATTTATAGATCCTCGTGATAGGCTAAACTACTCTACAAACA 1372
 Qy 1198 AFTCTGATGCTTTTGGTATTATGAGACTATACCTTTAACTGGAAAAGTAGAGATAATCACG 1257
 Db 1373 AFTCTGATGCTTTTGGTATTATGAGACTATACCTTTAACTGGAAAAGTAGAGATAATCACG 1432
 Qy 1258 ATGACACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCGAAGGAGAGAAATGCTTA 1317
 Db 1433 ATGACACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCGAAGGAGAGAAATGCTTA 1492
 Qy 1318 GCTATCATTTAGCCGGTGGT 1337
 Db 1493 GCTATCATTTAGCCTATGAT 1512

RESULT 9

AXX80492
 ID AAX80492 standard; cDNA; 1242 BP.
 XX AC
 AAX80492;
 XX
 17-OCT-2003 (revised)
 DT 26-AUG-1999 (first entry)
 XX
 Streptococcus equisimilis native streptokinase encoding cDNA.
 DE XX
 Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX
 Streptococcus dysgalactiae subsp. equisimilis.
 OS
 WO9931247-A1.
 PN
 24-JUN-1999.
 PD
 15-DEC-1998; 98WO-US026694.
 PF
 15-DEC-1997; 97US-0069497P.
 PR
 (HARD) HARVARD COLLEGE.
 XX
 Reed GL;
 FI
 WPI; 1999-395183/33.
 DR P-PSDB; AAY24794.
 XX
 N-terminally deleted streptokinase.
 PT
 Claim 44; Page 58-60; 73pp; English.
 PS
 XX

The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an expression vector comprising (1); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified

CC streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence encodes native streptokinase (nsk). (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;
 Query Match 69.28; Score 1148.6; DB 2; Length 1242;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 183 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCTCATCTGTCACACACAGCAATTTGTTGTT 242
 Db 1 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACACACAGCAATTTGTTGTT 60
 Qy 243 AGCGTTGCTGTAATGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATTC 302
 Db 61 AGCGTTGCTGTAATGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATTC 120
 Qy 303 GATCTAATCATCAAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 362
 Db 121 GATCTAATCATCAAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 180
 Qy 363 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
 Db 181 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 240
 Qy 423 AAGGCTATTCAAGAACCAATTTGATCGTAAACGTCCACAGTAAACGACGACTACTTTGAGTTC 482
 Db 241 AAGGCTATTCAAGAACCAATTTGATCGTAAACGTCCACAGTAAACGACGACTACTTTGAGTTC 300
 Qy 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTTACTTTGCTGAC 542
 Db 301 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTTACTTTGCTGAC 360
 Qy 543 AAAGATGTTTCGGTAACTTTGCCGACCCAACTCTGCAAGAAATTTTGTCTAAGCGGACAT 602
 Db 361 AAAGATGTTTCGGTAACTTTGCCGACCCAACTCTGCAAGAAATTTTGTCTAAGCGGACAT 420
 Qy 603 GTGCGGCTTAGACCCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTGTTGATGTG 662
 Db 421 GTGCGGCTTAGACCCATATAAAGAAAAACCAATAACAAACCAAGCGAAATCTGTTGATGTG 480
 Qy 663 GAATATATCTGTAAGTTTACTTCCCTTAAACCTGATGACGATTTTACAGACGAGTCTCAA 722
 Db 481 GAATATATCTGTAAGTTTACTTCCCTTAAACCTGATGACGATTTTACAGACGAGTCTCAA 540
 Qy 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTAATA 782
 Db 541 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTAATA 600
 Qy 783 GCTCAAGCAACAAAGCAATTTTAAACAAACCAAGCGCTATACGATTTATGAACGTCAC 842
 Db 601 GCTCAAGCAACAAAGCAATTTTAAACAAACCAAGCGCTATACGATTTATGAACGTCAC 660
 Qy 843 TCCTCAATCGTCACTCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 661 TCCTCAATCGTCACTCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
 Qy 903 TTTTACTTACCGTGTAAATATCGGAAACAGCTTATAGGATCAATAAATCTGTTCTG 962
 Db 721 TTTTACTTACCGTGTAAATATCGGAAACAGCTTATAGGATCAATAAATCTGTTCTG 780
 Qy 963 AATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 Db 781 AATGAAGAAATAAACAACACTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
 Qy 1023 GAAAGCCGATATGATCCCTTTGATCGCAGTCACTTTGAAACCTTTCAACATCAATACGTT 1082
 Db 841 GAAAGCCGATATGATCCCTTTGATCGCAGTCACTTTGAAACCTTTCAACATCAATACGTT 900
 Qy 1083 GATGTCGATACCAACGAATTTGCTTAAAGAGTGAGCAGCTCTTTAAGCAGCTAGCGAAGTAAAC 1142

Db 901 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAAC 960
 QY 1143 TTAGACTTCAGAGATTATATAGATCCCTCGTATAGGCTAAACTACTCTACAACTATCTC 1202
 Db 961 TTAGACTTCAGAGATTATATAGATCCCTCGTATAGGCTAAACTACTCTACAACTATCTC 1020
 QY 1203 GATGCTTTTGGTATATAGGACTATACCTTAACTCGAAAGTAGAGATAATCACGATGAC 1262
 Db 1021 GATGCTTTTGGTATATAGGACTATACCTTAACTCGAAAGTAGAGATAATCACGATGAC 1080
 QY 1263 ACCAAGCGTATACAGCTTTATATGGGCAAGCGACCGGAGAGAGAAATGCTAGCTAT 1322
 Db 1081 ACCAAGCGTATACAGCTTTATATGGGCAAGCGACCGGAGAGAGAAATGCTAGCTAT 1140
 QY 1323 CATTTAGCCGGTGGT 1337
 Db 1141 CATTTAGCCGTATGAT 1155

RESULT 10

AAA37633
 ID AAA37633 standard; DNA; 1245 BP.
 XX

AC AAA37633;

XX 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)

DE S. equimilis streptokinase coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibronectin; thrombolytic therapy;
 KW cardiovascular disorder; fibronectin; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers
 FH 1. 1245
 FT CDS /*tag= a
 FT /*product= "streptokinase"

XX EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99BP-00310541.

XX 24-DEC-1998; 98IN-DR003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sahní G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;

XX WPI: 2000-516032/47.

XX P-PSDB; AAY90282.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibronectin.

XX Example 3; Fig 3; 58pp; English.

XX This sequence represents the human Streptococcus equisimilis
 CC streptokinase coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibronectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable

CC animal or human Pg. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX

SQ Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match 69.2%; Score 1148.6; DB 3; Length 1245;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 183 ATTTGCTGGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAGAGCAATGGTTGT 242
 Db 1 ATTTGCTGGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAGAGCAATGGTTGT 60
 QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 120
 QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
 Db 121 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTGCATATAAATTGAGAAAGCTGACTACTA 422
 Db 181 AAACCATTTGCTACTGATAGTGGCGGATGTGCATATAAATTGAGAAAGCTGACTACTA 240
 QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTCAACGACGACTACTTTGAGGTC 482
 Db 241 AAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTCAACGACGACTACTTTGAGGTC 300
 QY 483 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 542
 Db 301 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 360
 QY 543 AAAGATGTTTCGGTAACTCTGCGGACCCAACTGTCGCAAGAAATTTTGTCAAGCGACAT 602
 Db 361 AAAGATGTTTCGGTAACTCTGCGGACCCAACTGTCGCAAGAAATTTTGTCAAGCGACAT 420
 QY 603 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGGAAATCTGTTGATGTG 662
 Db 421 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGGAAATCTGTTGATGTG 480
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCCAGACCGAGTCTCAAA 722
 Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCCAGACCGAGTCTCAAA 540
 QY 723 GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATCACAATCTCAAGAAATTTACTA 782
 Db 541 GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATCACAATCTCAAGAAATTTACTA 600
 QY 783 GCTCAAGCAACAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 601 GCTCAAGCAACAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
 QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
 Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
 QY 903 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 721 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGG 1022

Db 781 AATGAAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 Qy 1023 GAAGAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCAATACGTT 1082
 Db 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCAATACGTT 900
 Qy 1083 GATGTCGATACCAACAAATGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 901 GATGTCGATACCAACAAATGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
 Qy 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAACATCTC 1202
 Db 961 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAACATCTC 1020
 Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATATCAGGATGAC 1262
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATATCAGGATGAC 1080
 Qy 1263 ACCAACCGGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1322
 Db 1081 ACCAACCGGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1140
 Qy 1323 CATTTAGCCGCTGGT 1337
 Db 1141 CATTTAGCCCTATGAT 1155

RESULT 11
 ABA05546
 ID ABA05546 standard; cDNA; 1254 BP.
 XX
 AC ABA05546;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE Streptokinase cDNA.
 XX
 KW Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
 KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;
 KW gene therapy; maxadilan; ss.
 XX
 OS Unidentified.
 XX
 PN WO200185100-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US015209.
 XX
 PR 11-MAY-2000; 2000US-00569920.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Reddy VB, Lerner E;
 XX
 DR WPI; 2002-062184/08.
 XX
 PT New fusion protein or conjugate, useful for treating unstable angina,
 PT acute myocardial infarction or stroke, comprises a vasodilator
 PT polypeptide and a thrombolytic polypeptide, or active fragments of the
 PT polypeptides.
 XX
 PS Example 1; Fig 2; 37pp; English.
 XX
 CC The invention relates to a fusion protein or a conjugate comprising a
 CC vasodilator polypeptide, or its active fragment, and a thrombolytic
 CC polypeptide or its active fragment. The protein is useful for treating a
 CC subject suffering from a partially or totally occluded blood vessel,
 CC causing unstable angina, acute myocardial infarction or stroke. The
 CC polynucleotide encoding the polypeptide is useful in gene therapy. The
 CC vasodilatory action of the protein allows for the use of lower doses of a
 CC thrombolytic while maintaining the clot dissolving effectiveness of the

thrombolytic, and the use of lower doses of the thrombolytic reduces
 associated side effects. The present sequence is the streptokinase cDNA
 used in the construction of a Maxadilan-Streptokinase fusion protein.
 CC Maxadilan is a vasodilator peptide produced by the salivary gland of the
 CC New World sand fly
 XX
 SQ Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;
 Query Match 69.2%; Score 1148.6; DB 6; Length 1254;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 183 ATTCTCGACCTCGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTCTT 242
 Db 7 ATTCTCGACCTCGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTAGTCTT 66
 Qy 243 AGCGTTCTGTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
 Db 67 AGCGTTCTGTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 126
 Qy 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 362
 Db 127 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 186
 Qy 363 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
 Db 187 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 246
 Qy 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAAAGTCCACAGTAAACGACGACTCTTTGAGTGC 482
 Db 247 AAGGCTATTCAAGAACAAATTTGATCGCTTAAAGTCCACAGTAAACGACGACTCTTTGAGTGC 306
 Qy 483 ATTGATTTTGCAGCGATGCAACCAATTAAGTCAAGAACGCAAGGCTTACTTTGCTGAC 542
 Db 307 ATTGATTTTGCAGCGATGCAACCAATTAAGTCAAGAACGCAAGGCTTACTTTGCTGAC 366
 Qy 543 AAGATGTTTCGGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT 602
 Db 367 AAGATGTTTCGGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT 426
 Qy 603 GTGCGCGTTAGACCATATATAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 427 GTGCGCGTTAGACCATATATAAGAAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 486
 Qy 663 GAATATATCTGTA CAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCAGGCTCTAAA 722
 Db 487 GAATATATCTGTA CAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGCAGGCTCTAAA 546
 Qy 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 782
 Db 547 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 606
 Qy 783 GCTCAAGCAACAAAGCAATTTTAAACCAAAACCAAGCGCTATACGATTTTATGAACGTCAC 842
 Db 607 GCTCAAGCAACAAAGCAATTTTAAACCAAAACCAAGCGCTATACGATTTTATGAACGTCAC 666
 Qy 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 667 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 726
 Qy 903 TTTTACTTACCGTGTTTAAATTCGGGAACAGCTTATAGGATCAATAAAATCTGGTCTG 962
 Db 727 TTTTACTTACCGTGTTTAAATTCGGGAACAGCTTATAGGATCAATAAAATCTGGTCTG 786
 Qy 963 AATGAAGAAATAAACAACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 787 AATGAAGAAATAAACAACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 846
 Qy 1023 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCAATACGTT 1082
 Db 847 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCAATACGTT 906
 Qy 1083 GATGTCGATACCAACAAATGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142

Db 907 GAATGTCATACCAAGCAATTCCTAAAGTAGGAGCTCTTAAACAGCTAGCGAAGCTAAC 966
 QY 1143 TTAGACTTCAGAGATTTATAGGATCCCTCGTGATAAGGCTAAACTACTCTACAAATCTC 1202
 Db 967 TTAGACTTCAGAGATTTATAGGATCCCTCGTGATAAGGCTAAACTACTCTACAAATCTC 1026
 QY 1203 GATGCTTTTGGTATATATGGACTATACCTTAACTGCGAAGTAGAGATAATCACGATGAC 1262
 Db 1027 GATGCTTTTGGTATATATGGACTATACCTTAACTGCGAAGTAGAGATAATCACGATGAC 1086
 QY 1263 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGAGAGAGAGAAATGCTAGCTAT 1322
 Db 1087 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGAGAGAGAGAAATGCTAGCTAT 1146
 QY 1323 CATTAGCCGCTGGT 1337
 Db 1147 CATTAGCCCTATGAT 1161

RESULT 12

ABA05547

ID ABA05547 standard; DNA; 8893 BP.

XX ABA05547;

AC ABA05547;

XX 26-FEB-2002 (first entry)

XX Maxadilan-streptokinase fusion protein plasmid pTVB3maxstk.

XX Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;

KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;

KW gene therapy; maxadilan; sand fly; plasmid; ds.

XX Lutzomyia longipalpis.

OS Unidentified.

OS Synthetic.

OS Chimeric.

XX WO200185100-A2.

XX 15-NOV-2001.

XX 10-MAY-2001; 2001WO-US015209.

XX 11-MAY-2000; 2000US-00569920.

XX (GEO) GEN HOSPITAL CORP.

XX Reddy VB, Lerner E;

XX WPI; 2002-062184/08.

XX New fusion protein or conjugate, useful for treating unstable angina,

XX acute myocardial infarction or stroke, comprises a vasodilator

XX polypeptide and a thrombolytic polypeptide, or active fragments of the

XX polypeptides.

XX Example 1; Fig 3; 37pp; English.

XX The invention relates to a fusion protein or a conjugate comprising a

XX vasodilator polypeptide, or its active fragment, and a thrombolytic

XX polypeptide or its active fragment. The protein is useful for treating a

XX subject suffering from a partially or totally occluded blood vessel,

XX causing unstable angina, acute myocardial infarction or stroke. The

XX polynucleotide encoding the polypeptide is useful in gene therapy. The

XX vasodilatory action of the protein allows for the use of lower doses of a

XX thrombolytic while maintaining the clot dissolving effectiveness of the

XX thrombolytic, and the use of lower doses of the thrombolytic reduces

XX associated side effects. The present sequence is the plasmid used for the

XX preparation of Maxadilan-Streptokinase fusion protein. It contains

XX maxadilan cDNA, which encodes a vasodilator peptide produced by the

XX salivary gland of the New World sand fly, and streptokinase cDNA

XX

SQ Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match 69.2%; Score 1148.6; DB 6; Length 8893;

Best Local Similarity 99.7%; Pred. No. 3.3e-314;

Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY

183 ATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGTT 242

Db

5935 ATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGTT 5994

QY

243 AGCGTTGCTGGTACTGTTGAGGGGCGAATCAAGACATTGCTTTAAATTTTGAATTC 302

Db

5995 AGCGTTGCTGGTACTGTTGAGGGGCGAATCAAGACATTGCTTTAAATTTTGAATTC 6054

QY

303 GATCTAAACATCACGACCTGCTCATGAGGGAAGACAGCAGGCTTAAGTCCAAATCA 362

Db

6055 GATCTAAACATCACGACCTGCTCATGAGGGAAGACAGCAGGCTTAAGTCCAAATCA 6114

QY

363 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTTA 422

Db

6115 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTTA 6174

QY

423 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTTCCACAGTAAACGACGACTTCTTGGGTC 482

Db

6175 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTTCCACAGTAAACGACGACTTCTTGGGTC 6234

QY

483 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGTAAACGTTCCACAGTAAACGACGACTTCTTGGGTC 542

Db

6235 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGTAAACGTTCCACAGTAAACGACGACTTCTTGGGTC 6294

QY

543 AAAGATGTTTCGGTAAACCTTGCAGACCCCAACCTGTCGAAGAAATTTTTCGTAAGCGGACAT 602

Db

6295 AAAGATGTTTCGGTAAACCTTGCAGACCCCAACCTGTCGAAGAAATTTTTCGTAAGCGGACAT 6354

QY

603 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGGAAATCTGTTGATGTCG 662

Db

6355 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGGAAATCTGTTGATGTCG 6414

QY

663 GAATATACGTACAGTTTATCTCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTCTCAAA 722

Db

6415 GAATATACGTACAGTTTATCTCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTCTCAAA 6474

QY

723 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACACATCAGATCTCAAGAAATTACTA 782

Db

6475 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACACATCAGATCTCAAGAAATTACTA 6534

QY

783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 842

Db

6535 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTCAC 6594

QY

843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902

Db

6595 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 6654

QY

903 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962

Db

6655 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 6714

QY

963 AATCAAGAAATAAACAACACACTGCTGATCTCTGGAATATATTAGTCTCTTAAAAAAGGG 1022

Db

6715 AATCAAGAAATAAACAACACACTGCTGATCTCTGGAATATATTAGTCTCTTAAAAAAGGG 6774

QY

1023 GAAAGCGGTATGATCCCTTTGATCGGAGTCACTTTGAAATCTGTTCAACCATCAATATACGTT 1082

Db

6775 GAAAGCGGTATGATCCCTTTGATCGGAGTCACTTTGAAATCTGTTCAACCATCAATATACGTT 6834

QY

1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAAGAGCTCTTAAACAGCTAGCGAAGCTAAC 1142

Db

6835 GATGTCGATACCAACGAATTTGCTAAAAAGTGAAGAGCTCTTAAACAGCTAGCGAAGCTAAC 6894

QY

1143 TTAGACTTCAGAGATTTATAGGATCCCTCGTGATAAGGCTAAAGTCTCTTCAACATCTC 1202

Db 6895 TTAGACTTCAGAGATTATATACGATCCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 6954

Qy 1203 GATGCTTTTGGTATTATGGACTACTACCTTAAGTGGAAAAGTAGAGGATAATCAGATGAC 1262

Db 6955 GATGCTTTTGGTATTATGGACTACTACCTTAAGTGGAAAAGTAGAGGATAATCAGATGAC 7014

Qy 1263 ACCAACCGTATCATACCGGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1322

Db 7015 ACCAACCGTATCATACCGGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 7074

Qy 1323 CATTTAGCCGGTGGT 1337

Db 7075 CATTTAGCCCTATGAT 7089

RESULT 13

AXX16632

ID AAX16632 standard; DNA; 1242 BP.

AC AAX16632;

XX 17-OCT-2003 (revised)

DT 04-MAY-1999 (first entry)

XX Streptococcus equisimilis native streptokinase encoding DNA.

XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;

KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;

KW serine protease; fibrin; blood clot; thrombolytic;

KW vascular thromboembolytic symptom; acute myocardial infarction;

KW fibrinolysis; resistance; ds.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS

XX

Key Location/Qualifiers

FT CDS 1..1242

FT /tag= a

FT /transl_except= (pos:40..42,aa:Asn)

FT /note= "no stop codon given"

XX

PN US5876999-A.

XX

PD 02-MAR-1999.

XX

PF 06-DEC-1995; 95US-00568393.

XX

PR 06-DEC-1995; 95US-00568393.

XX

PA (NASC-) NAT SCI COUNCIL.

XX

PI Wu H;

XX

DR WPI; 1999-189643/16.

DR P-PSDB; AAW94664.

XX

PT Mutant streptokinase polypeptide - useful as plasmin-resistant

PT thrombolytic agent.

XX

PS Claim 1; Col 7-10; 17pp; English.

XX

CC The present invention describes a mutant streptokinase (SK) polypeptide

CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61

CC segment of the corresponding native SK is replaced by another amino acid.

CC The present sequence encodes native SK. SK is a secretory protein of

CC haemolytic Streptococcus able to activate human plasminogen (HPlg) to

CC plasmin (HPlm), which is a serine protease able to catalyse the

CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic

CC agent in the treatment of vascular thromboembolytic symptoms such as

CC acute myocardial infarction. Compared with wild-type SK, the X598 mutant

CC is more resistant to degradation by human plasmin and is more effective

CC both in acting as a fibrinolytic agent and in activating human plasminogen.

CC (Updated on 17-OCT-2003 to standardise OS field)

XX

SQ Sequence 1242 BP; 423 A; 268 C; 237 G; 314 T; 0 U; 0 Other;

Query Match 69.1%; Score 1147; DB 2; Length 1242;

Best Local Similarity 99.6%; Pred. No. 3.3e-314;

Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 183 ATTCTCGACCTCGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGTTGTTT 242

Db 1 ATTCTCGACCTCGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGTTGTTT 60

Qy 243 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302

Db 61 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120

Qy 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 362

Db 121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 180

Qy 363 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 422

Db 181 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 240

Qy 423 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGACGACTCTTTGAGTGC 482

Db 241 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGACGACTCTTTGAGTGC 300

Qy 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTACTTTGCTGAC 542

Db 301 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTACTTTGCTGAC 360

Qy 543 AAAGATGGTTCGGTAACTCTTCCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 602

Db 361 AAAGATGGTTCGGTAACTCTTCCCGACCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 420

Qy 603 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGANTG 662

Db 421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGANTG 480

Qy 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACGAGGCTCAAA 722

Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACGAGGCTCAAA 540

Qy 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782

Db 541 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600

Qy 783 GCTCAAGCAACAAGCATTTTAAACAAAAACCAAGCGGCTATACGATTTATGAACGTGAC 842

Db 601 GCTCAAGCAACAAGCATTTTAAACAAAAACCAAGCGGCTATACGATTTATGAACGTGAC 660

Qy 843 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902

Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720

Qy 903 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962

Db 721 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780

Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 1022

Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 840

Qy 1023 GAAAGCCGTAATGATCCCTTTGATCGGAGTCACTTTGAAACTGTTTACCATCAAAATACGTT 1082

Db 841 GAAAGCCGTAATGATCCCTTTGATCGGAGTCACTTTGAAACTGTTTACCATCAAAATACGTT 900

Qy 1083 GATGTCATACCAACGAAATGCTTAAAGTAGGAGAGCTCTTAAAGCTAGCGAGCTAAC 1142

Db 901 GATGTCATACCAACGAAATGCTTAAAGTAGGAGAGCTCTTAAAGCTAGCGAGCTAAC 960

Qy 1143 TTAGACTTTCAGAGATTTTATACGATCCCTCGTGAATAGGCTAAACTACTCTACAAATCTC 1202

Db 961 TTAGACTTTCAGAGATTTTATACGATCCCTCGTGAATAGGCTAAACTACTCTACAAATCTC 1020

```
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1262
Db |||||||
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT 1322
Db |||||||
QY 1323 CATTTAGCCGGTGGT 1337
Db |||||||
QY 1141 CATTTAGCCCTATGAT 1155
Db |||||||

RESULT 14
AA16633
ID AAX16633 standard; DNA; 1242 BP.
AC AAX16633;
XX
04-MAY-1999 (first entry)
Streptococcus equisimilis mutant streptokinase K59B encoding DNA.
KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance; ds.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
FH
FX
FT CDS
FT 1..1242
FT /*tag= a
FT /transl_except= (pos:40..42,aa:Asn)
FT /note= "no stop codon given"
XX
PN US5876999-A.
XX
XX
PD 02-MAR-1999.
XX
XX
PF 06-DEC-1995; 95US-00568393.
XX
XX
PR 06-DEC-1995; 95US-00568393.
XX
XX
PA (NASC-) NAT SCI COUNCIL.
XX
XX
PI Wu H;
XX
XX
DR WPI; 1999-189643/16.
DR P-PSDB; AAX94665.
XX
XX
PT Mutant streptokinase polypeptide - useful as plasmin-resistant
PT thrombolytic agent.
XX
XX
PS Claim 4; Col 11-14; 17pp; English.
XX
CC The present invention describes a mutant streptokinase (SK) polypeptide
CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
CC segment of the corresponding native SK is replaced by another amino acid.
CC The present sequence encodes mutant SK K59E. SK is a secretory protein of
CC haemolytic streptococcus able to activate human plasminogen (HPlg) to
CC plasmin (HPlm), which is a serine protease able to catalyse the
CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
CC agent in the treatment of vascular thromboembolytic symptoms such as
CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
CC is more resistant to degradation by human plasmin and is more effective
CC both in acting as a fibrinolytic agent and in activating human plasminogen
XX
SQ Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;
```

```
Query Match 68.9%; Score 1143.8; DB 2; Length 1242;
Best Local Similarity 99.4%; Pred. No. 2e-313;
Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCCCAATGGTGT 242
Db |||||||
QY 1 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCCCAATGGTGT 60
Db |||||||
QY 243 AGCGTGTGCTGTCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db |||||||
QY 61 AGCGTGTGCTGTCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
Db |||||||
QY 303 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
Db |||||||
QY 121 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCGGAATCA 180
Db |||||||
QY 363 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTACTA 422
Db |||||||
QY 181 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTACTA 240
Db |||||||
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAAACGACGACTACTTTGAGGTC 482
Db |||||||
QY 241 AAGGCTATTCAAGAACAAATTGATCGCTAACCGTCCACAGTAAACGACGACTACTTTGAGGTC 300
Db |||||||
QY 483 ATTGATTTTGAAGCGATGCAACCACTTACTGATGAAACGGCAAGGTCTACTTTGCTGAC 542
Db |||||||
QY 301 ATTGATTTTGAAGCGATGCAACCACTTACTGATGAAACGGCAAGGTCTACTTTGCTGAC 360
Db |||||||
QY 543 AAAGATGTTTCGGTAAACCTTGGCGGACCCCACTGTCGCAAGAAATTTTGTCTAAGCGGACAT 602
Db |||||||
QY 361 AAAGATGTTTCGGTAAACCTTGGCGGACCCCACTGTCGCAAGAAATTTTGTCTAAGCGGACAT 420
Db |||||||
QY 603 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAATCTGTTGATGTCG 662
Db |||||||
QY 421 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAATCTGTTGATGTCG 480
Db |||||||
QY 663 GAATATAGTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCCAGACAGGTCTCAAA 722
Db |||||||
QY 481 GAATATAGTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCCAGACAGGTCTCAAA 540
Db |||||||
QY 723 GATCTAAGCTATTGAAACCACTAGCTATCGGTGACCACTCACATCTCAAGAAATTA 782
Db |||||||
QY 541 GATCTAAGCTATTGAAACCACTAGCTATCGGTGACCACTCACATCTCAAGAAATTA 600
Db |||||||
QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTATGAAGTAC 842
Db |||||||
QY 601 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTATGAACGTCAC 660
Db |||||||
QY 843 TCCTCAATCGTCACCTCATGACAATGACATTTTCCGTAGCATTTTACCAGTCAATGATCAAGAG 902
Db |||||||
QY 661 TCCTCAATCGTCACCTCATGACAATGACATTTTCCGTAGCATTTTACCAGTCAATGATCAAGAG 720
Db |||||||
QY 903 TTTTACTTACCGTGTAAAAATCGGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db |||||||
QY 721 TTTTACTTACCGTGTAAAAATCGGGAAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
Db |||||||
QY 963 AATCAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
Db |||||||
QY 781 AATCAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 840
Db |||||||
QY 1023 GAAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACATCAATACGTT 1082
Db |||||||
QY 841 GAAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACATCAATACGTT 900
Db |||||||
QY 1083 GATCTCGATACCAAGCAATTTCTAAAAAGTCAAGAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
Db |||||||
QY 901 GATCTCGATACCAAGCAATTTCTAAAAAGTCAAGAGCTCTTAAACAGCTAGCGAAGCTAAC 960
Db |||||||
QY 1143 TTAGACTTCAGAGATTTTATACGATCCCTGCTGATAGGCTAAACTACTCTACAACTCTC 1202
Db |||||||
QY 961 TTAGACTTCAGAGATTTTATACGATCCCTGCTGATAGGCTAAACTACTCTACAACTCTC 1020
Db |||||||
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
```


Db 3093 TCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAA 3152
Qy 1313 TGCTAGCTATCATTTAGCCGTTGGT 1337
Db 3153 TGCTAGCTATCATTTAGCCTATGAT 3177

Search completed: January 31, 2006, 18:17:50
Job time : 1005.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:24:19 ; Search time 293.256 Seconds
(without alignments)
10068.109 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaacccgcagctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/H COMB.seq:*
6: /cgn2_6/prodata/1/ina/PTUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfilee1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150	69.2	2385	3	US-09-211-542A-1
2	1148.6	69.2	1242	3	US-09-211-542A-5
3	1147	69.1	1242	2	US-08-568-393B-1
4	1143.8	68.9	1242	2	US-08-568-393B-2
5	1134.6	68.3	2566	2	US-08-488-940-19
6	1127.8	67.9	1245	2	US-07-703-778D-1
7	1127.6	67.9	1257	2	US-07-854-596B-25
8	1126.2	67.8	1512	2	US-07-854-596B-27
9	1124.6	67.7	1317	2	US-07-854-596B-18
10	1124.6	67.7	1335	2	US-07-854-596B-14
11	1124.6	67.7	1458	2	US-07-854-596B-42
12	1124.6	67.7	1467	2	US-07-854-596B-46
13	1124.6	67.7	2589	2	US-07-854-596B-34
14	1096.6	66.0	1122	3	US-09-374-038-13
15	1096.6	66.0	1122	3	US-09-558-179-13
16	1096.6	66.0	1158	3	US-09-374-038-14
17	1096.6	66.0	1158	3	US-09-558-179-14
18	1089.2	65.6	1209	3	US-09-374-038-11
19	1089.2	65.6	1209	3	US-09-558-179-11
20	1089.2	65.6	1245	3	US-09-374-038-12
21	1089.2	65.6	1245	3	US-09-558-179-12
22	1080.4	65.0	2253	2	US-07-854-596B-39
23	1078.8	64.9	1119	2	US-07-854-596B-30
24	976.2	58.8	1068	3	US-09-211-542A-11

25	973.2	58.6	2208	3	US-09-211-542A-3	Sequence 3, Appli
26	762.8	45.9	1262	9	5240845-3	Patent No. 5240845
27	757.8	45.6	1242	9	5240845-2	Patent No. 5240845
28	450	27.1	450	3	US-09-211-542A-13	Sequence 13, Appl
29	316.8	19.1	7679	3	US-09-220-132-38	Sequence 38, Appl
30	316.8	19.1	7680	3	US-09-023-655-1289	Sequence 1289, Ap
31	316.8	19.1	7680	6	PCT-US95-09819-6	Sequence 6, Appli
32	316.8	19.1	7705	2	US-08-259-569-16	Sequence 16, Appl
33	316.8	19.1	7705	9	US-08-826-885-16	Sequence 16, Appl
34	316.8	19.1	7705	9	5455158-2	Patent No. 5455158
35	316.8	19.1	7803	2	US-08-551-356-1	Sequence 1, Appli
36	316.8	19.1	7803	6	PCT-US93-12887-1	Sequence 1, Appli
37	310.4	18.7	8044	3	US-09-566-921-135	Sequence 135, App
38	178	10.7	4100	3	US-09-813-718-5	Sequence 5, Appli
39	178	10.7	4682	3	US-09-813-718-3	Sequence 3, Appli
40	178	10.7	4682	3	US-09-813-718-7	Sequence 7, Appli
41	178	10.7	4742	3	US-09-813-718-15	Sequence 15, Appl
42	178	10.7	4811	3	US-09-813-718-13	Sequence 13, Appl
43	178	10.7	4877	3	US-09-813-718-11	Sequence 11, Appl
44	178	10.7	5018	3	US-09-813-718-9	Sequence 9, Appli
45	178	10.7	5174	3	US-09-813-718-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-211-542A-1
; Sequence 1, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2385
US-09-211-542A-1

Query Match 69.2%; Score 1150; DB 3; Length 2385;

Best Local Similarity 99.6%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 180 ATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCCAATGGTT 239
DB 1141 AGGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCCAATGGTT 1200

QY 240 GTTAGCGTTGCTGCTAGTGTGAGGGACGAATCAAGACATAGTCTTAAATTTTGA 299
DB 1201 GTTAGCGTTGCTGCTAGTGTGAGGGACGAATCAAGACATAGTCTTAAATTTTGA 1260

QY 300 ATCGATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAA 359
DB 1261 ATCGATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAA 1320

QY 360 TCAAAACCAATTTGCTACTGATAGTGGCGCATGTCAATAAATCTTGAGAAAGCTGACTTA 419
DB 1321 TCAAAACCAATTTGCTACTGATAGTGGCGCATGTCAATAAATCTTGAGAAAGCTGACTTA 1380

QY 420 CTAAGGCTATTCAAGACCAATTTGATCGCTTAAGTCCACAGTAAACGACGACTACTTTGAG 479
DB 1381 CTAAGGCTATTCAAGACCAATTTGATCGCTTAAGTCCACAGTAAACGACGACTACTTTGAG 1440

QY 480 GTCAATGATTTTCAAGCGATGCAACCAATTTACTGATCGAAACGCGCAAGTCTACTTTGCT 539
DB 1441 GTCAATGATTTTCAAGCGATGCAACCAATTTACTGATCGAAACGCGCAAGTCTACTTTGCT 1500

QY 540 GACAAAGATGTTTGGTACCTTTGCGGACCCAACTGCTCAAGAAATTTTGTGTAAGCGGA 599
DB 1501 GACAAAGATGTTTGGTACCTTTGCGGACCCAACTGCTCAAGAAATTTTGTGTAAGCGGA 1560

QY 600 CATGTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAATCTGTTGAT 659
DB 1561 CATGTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAATCTGTTGAT 1620

QY 660 GTGGAATATCTGTCAGATTTTACTCCCTTAAACCCCTGATCAGCATTTTACAGCAGGTCTC 719
DB 1621 GTGGAATATCTGTCAGATTTTACTCCCTTAAACCCCTGATCAGCATTTTACAGCAGGTCTC 1680

QY 720 AAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 779
DB 1681 AAGATATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1740

QY 780 CTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATAGCATTTATGAACGT 839
DB 1741 CTAGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATAGCATTTATGAACGT 1800

QY 840 GACTCTCAATCTGCTACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAA 899
DB 1801 GACTCTCAATCTGCTACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAA 1860

QY 900 GAGTTTACTTACCGTGTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGT 959
DB 1861 GAGTTTACTTACCGTGTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGT 1920

QY 960 CTGAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTAAGTCTTTAAAAA 1019
DB 1921 CTGAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTAAGTCTTTAAAAA 1980

QY 1020 GGGGAAAGCCGATGATCCCTTTGATCGGAGTCACTTTGAAATCTGTTACCAATCAATAC 1079
DB 1981 GGGGAAAGCCGATGATCCCTTTGATCGGAGTCACTTTGAAATCTGTTACCAATCAATAC 2040

QY 1080 GTTGATGTGATACCAACCAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGGAAAGCT 1139
DB 2041 GTTGATGTGATACCAACCAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGGAAAGCT 2100

QY 1140 AACTTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTAAATCTACTCTCAACAAT 1199
DB 2101 AACTTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTAAATCTACTCTCAACAAT 2160

QY 1200 CTCGATGCTTTTGGTATTGATGATACCTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1259

DB 2161 CTCGATGCTTTTGGTATTATGGACTATACCTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2220

QY 1260 GACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGTAAAGTAAAGTAAAGT 1319
DB 2221 GACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGTAAAGTAAAGTAAAGT 2280

QY 1320 TATCATTTAGCCGGTGGT 1337
DB 2281 TATCATTTAGCCCTATGAT 2298

RESULT 2
US-09-211-542A-5
; Sequence 5, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/211,542A
; APPLICATION NUMBER: 15-December-1998
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1242
; US-09-211-542A-5

Query Match 69.2%; Score 1148.6; DB 3; Length 1242;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACACGCAATTTGGTTGTT 242
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACACGCAATTTAGTTGTT 60

QY 243 AGCGTCTGCTACTGTTGAGGGACCAATCAACACATTAAGTCTTAAATTTTGAATC 302
DB 61 AGCGTCTGCTACTGTTGAGGGACCAATCAACACATTAAGTCTTAAATTTTGAATC 120

QY 303 GATCTAAACATCACCGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAAATCA 362
DB 121 GATCTAAACATCACCGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAAATCA 180

```
QY 363 AAACCAATTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 181 AAACCAATTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
QY 423 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGAGCACTACTTTGAGTGC 482
Db 241 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGAGCACTACTTTGAGTGC 300
QY 483 ATTGATTTTGCAGCGATGCAACCAATTACTGTATCGAAACGGCAAGTCTACTTTGCTGAC 542
Db 301 ATTGATTTTGCAGCGATGCAACCAATTACTGTATCGAAACGGCAAGTCTACTTTGCTGAC 360
QY 543 AAAGATGGTTGGTAACTTCCGCAACCACTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 361 AAAGATGGTTGGTAACTTCCGCAACCACTGTCCAAGAAATTTTGTCTAAGCGGACAT 420
QY 603 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 662
Db 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 480
QY 663 GAATATACTCTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 722
Db 481 GAATATACTCTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 540
QY 723 GATATAGCTATTGAAACCACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 782
Db 541 GATATAGCTATTGAAACCACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 600
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 842
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTGAC 660
QY 843 TCCCTCAATCTGCTACTGATCAATGACATTTCCGTACGATTTTACCAATGATCAAGAG 902
Db 661 TCCCTCAATCTGCTACTGATCAATGACATTTCCGTACGATTTTACCAATGATCAAGAG 720
QY 903 TTTACTTACCGTGTAAATAACCGGAACCAAGCTTATAGGATCAATAAAATCTGGTCTG 962
Db 721 TTTACTTACCGTGTAAATAACCGGAACCAAGCTTATAGGATCAATAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACCACTGACCTGATCTTGAGAAATATTACGCTCTTAAAAAAGGG 1022
Db 781 AATGAAGAAATAAACCACTGACCTGATCTTGAGAAATATTACGCTCTTAAAAAAGGG 840
QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTTCAACCATCAATACGTT 1082
Db 841 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTTCAACCATCAATACGTT 900
QY 1083 GATGTCGATACCAACGAATTTGCTAAAGTGAAGTGAAGCTCTTAAACAGCTAGCGAAGTAA 1142
Db 901 GATGTCGATACCAACGAATTTGCTAAAGTGAAGTGAAGCTCTTAAACAGCTAGCGAAGTAA 960
QY 1143 TTAGACTTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTTCTTACAAACATCTC 1202
Db 961 TTAGACTTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTTCTTACAAACATCTC 1020
QY 1203 GATGCTTTTGGTATTATGATATACCTTAACTGGAAGTGAAGGATAATCACGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGATATACCTTAACTGGAAGTGAAGGATAATCACGATGAC 1080
QY 1263 ACCAACGGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGAGATGCTAGCTAT 1322
Db 1081 ACCAACGGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGAGATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCCTATGAT 1155
```

RESULT 3
US-08-568-393B-1
; Sequence 1, Application US/08568393B
; Patent No. 5876999

```
; GENERAL INFORMATION:  
; APPLICANT: Hua-Lin Wu  
; APPLICANT: Guey-Yueh Shi  
; TITLE OF INVENTION: Preparation of novel streptokinase  
; TITLE OF INVENTION: mutants as improved thrombolytic agents  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jeing & Chang  
; STREET: Two No. 5876999th Second Street, Suite 290  
; CITY: San Jose  
; STATE: California  
; COUNTRY: USA  
; ZIP: 95113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB  
; MEDIUM TYPE: storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 6.1 on Window 3.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/568,393B  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Chi-Ping Chang  
; REGISTRATION NUMBER: 37,798  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (408) 288-8585  
; TELEFAX: (408) 288-8386  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1242 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL: N  
; ANTI-SENSE: N  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus equisimilis H46A  
; INDIVIDUAL ISOLATE: Maik, H., Roe, B., and Ferretti, J. J.;  
; INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A" from Gene 34:357-362 (1985).  
; CELL TYPE: Streptococcus equisimilis H46A  
; US-08-568-393B-1
```

```
Query Match 69.1%; Score 1147; DB 2; Length 1242;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATTTGTTGTT 242  
Db 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATTTGTTGTT 60  
QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302  
Db 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 120  
QY 303 GATCTAACATCACACCTGCTCTGAGGAAAGACAGAGCAAGCTTAAGTCCAAATCA 362  
Db 121 GATCTAACATCACACCTGCTCTGAGGAAAGACAGAGCAAGCTTAAGTCCAAATCA 180  
QY 363 AAACCAATTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422  
Db 181 AAACCAATTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240  
QY 423 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGAGCACTACTTTGAGTGC 482  
Db 241 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTAACGAGCACTACTTTGAGTGC 300  
QY 483 ATTGATTTTGCAGCGATGCAACCAATTACTGTATCGAAACGGCAAGTCTACTTTGCTGAC 542
```

```

Db 301 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 360
QY 543 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
Db 361 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 420
QY 603 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 421 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY 663 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGTCTCAAA 722
Db 481 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGTCTCAAA 540
QY 723 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATCTACTA 782
Db 541 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATCTACTA 600
QY 783 GCTCAAGCAACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 601 GCTCAAGCAACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 843 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCAAGAG 902
Db 661 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCAAGAG 720
QY 903 TTTTACTACCGTGTAAATAATCGGACCAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 721 TTTTACTACCGTGTAAATAATCGGACCAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
QY 1023 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTGAAACTGTTTCAACATCAAAATACGTT 1082
Db 841 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTGAAACTGTTTCAACATCAAAATACGTT 900
QY 1083 GATGTCGATACCAAGAAATGCTTAAAGTGACAGCTCTTAAACAGTACGGAACGTAAC 1142
Db 901 GATGTCGATACCAAGAAATGCTTAAAGTGACAGCTCTTAAACAGTACGGAACGTAAC 960
QY 1143 TTAGACTTCCAGAGATTTATACGATCTCTGTTGATAAGCTAAACTTCTTCAACAATCTC 1202
Db 961 TTAGACTTCCAGAGATTTATACGATCTCTGTTGATAAGCTAAACTTCTTCAACAATCTC 1020
QY 1203 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATATCAGGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGGATATCAGGATGAC 1080
QY 1263 ACCAACCGTATCATTAACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
Db 1081 ACCAACCGTATCATTAACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCCTATGAT 1155

```

RESULT 4

```

US-08-568-393B-2
; Sequence 2, Application US/08568393B
; Patent No. 5876999
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose

```

```

; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408)288-8386
; TELEFAX: (408)288-8386
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: SK-K59E
; LOCATION: DNA sequence No. 5876999174 and 175 have been changed
; LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
; LOCATION: from lys to Glu.
; OTHER INFORMATION:
; US-08-568-393B-2

```

Query Match 68.9%; Score 1143.8; DB 2; Length 1242;

Best Local Similarity 99.4%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATGCTGTTGTT 242
Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATGCTGTTGTT 60
QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 120
QY 303 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAATCA 362
Db 121 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAATCA 180
QY 363 AAACCATTTGCTACTGATAGTGGCGCATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 181 AAACCATTTGCTACTGATAGTGGCGCATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
QY 423 AAGGCTATTCAAGAAACAAATTTGATCGCTTAAACGTCACAGTAACGACGACTACTTTGAGTGC 482
Db 241 AAGGCTATTCAAGAAACAAATTTGATCGCTTAAACGTCACAGTAACGACGACTACTTTGAGTGC 300
QY 483 ATTGATTTTGAAGCGATGCAACCAATTAAGTCAAGAAACGCAAGGCTTACTTTGCTGAC 542
Db 301 ATTGATTTTGAAGCGATGCAACCAATTAAGTCAAGAAACGCAAGGCTTACTTTGCTGAC 360
QY 543 AAAGATGGTTCGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
Db 361 AAAGATGGTTCGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 420
QY 603 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 662
Db 421 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 480

```

QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGATTTTCAGACCGAGTCTCAAA 722
Db 481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGATTTTCAGACCGAGTCTCAAA 540
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 782
Db 541 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600
QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCCACCGAGCTATACGATTTATGAACGTGAC 842
Db 601 GCTCAAGCACAAGCAATTTTAAACCAAAACCCACCGAGCTATACGATTTATGAACGTGAC 660
QY 843 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGAGATTTTACCAATGATCAAGAG 902
Db 661 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGAGATTTTACCAATGATCAAGAG 720
QY 903 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCTGGTCTG 962
Db 721 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 781 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
QY 1023 GAAAAGCCGTATGATCCCTTTGTGTCGAGTCACTTGAACCTGTTTCAACATCAAAATACGTT 1082
Db 841 GAAAAGCCGTATGATCCCTTTGTGTCGAGTCACTTGAACCTGTTTCAACATCAAAATACGTT 900
QY 1083 GATGTCGATACCAACGAATTTGCTTAAAGATGAGGAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
Db 901 GATGTCGATACCAACGAATTTGCTTAAAGATGAGGAGCTCTTAAACAGCTAGCGAAGCTAAC 960
QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAACTACTCTCAACAATCTC 1202
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAACTACTCTCAACAATCTC 1020
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGGATGAC 1080
QY 1263 ACCAAGCTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1322
Db 1081 ACCAAGCTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCCTATGAT 1155

RESULT 5
US-08-488-940-19
; Sequence 19, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-488-940-19

Query Match 68.3%; Score 1134.6; DB 2; Length 2566;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1149; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGGTTGTT 242
Db 897 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGGTTGTT 956
QY 243 AGCGTTGCTGGTACTGTTGAGGGGAGGAATCAAGACATTTAGTCTTAAATTTTTCGAAATC 302
Db 957 AGCGTTGCTGGTACTGTTGAGGGGAGGAATCAAGACATTTAGTCTTAAATTTTTCGAAATC 1016
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAAATCA 362
Db 1017 GATCTAACATCACGACCTGCTCAT--AGGAAAGACAGACGAGGCTTAAGTCCAAAATCA 1074
QY 363 AAACCATTTGCTACTGATGAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 1075 AAACCATTTGCTACTGATGAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 1134
QY 423 AAGGCTATTCAAGACCAATTCGCTAAGCTCCACAGTCCACAGTACGACGACTCTTTCAGGTC 482
Db 1135 AAGGCTATTCAAGACCAATTCGCTAAGCTCCACAGTCCACAGTACGACGACTCTTTCAGGTC 1194
QY 483 ATTGATTTTGAACGATGCAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
Db 1195 ATTGATTTTGAACGATGCAACCAATTAATCTGATCGAAACGGCAAGGTCTACTTTGCTGAC 1254
QY 543 AAAGATGGTTCGGTAAACCTTGCCTGACCCCAACCTGTCGAAGAAATTTTTCCTAAGCGGACAT 602
Db 1255 AAAGATGGTTCGGTAAACCTTGCCTGACCCCAACCTGTCGAAGAAATTTTTCCTAAGCGGACAT 1314
QY 603 GTGCGCGTTAGACCATATATAAGAAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 1315 GTGCGCGTTAGACCATATATAAGAAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 1374
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGATTTTCAGACCGAGTCTCAAA 722
Db 1375 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGATTTTCAGACCGAGTCTCAAA 1434
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 782
Db 1435 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1494
QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCCACCGAGCTATACGATTTATGAACGTGAC 842
Db 1495 GCTCAAGCACAAGCAATTTTAAACCAAAACCCACCGAGCTATACGATTTATGAACGTGAC 1554
QY 843 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGAGATTTTACCAATGATCAAGAG 902
Db 1555 TCCTCAATCGTCACTGATGACATGACATTTTCCGTAGAGATTTTACCAATGATCAAGAG 1614
QY 903 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCTGGTCTG 962
Db 1615 TTTTACTTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATATAAATAATCTGGTCTG 1674

```

QY 963 AATGAAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 1675 AATGAAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1734
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATACGTT 1082
Db 1735 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATACGTT 1794
QY 1083 GATGTCGATACCAACGAATTTGCTTAAAAAGTGACGCTCTTAAACGCTAGCGAACGTAAC 1142
Db 1795 GATGTCGATACCAACGAATTTGCTTAAAAAGTGACGCTCTTAAACGCTAGCGAACGTAAC 1854
QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATGATGATGATGATGATGATGATGAT 1202
Db 1855 TTAGACTTCAGAGATTTATACGATCTCTGATGATGATGATGATGATGATGATGATGAT 1914
QY 1203 GATGCTTTTGGTATTTAGGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1262
Db 1915 GATGCTTTTGGTATTTAGGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1974
QY 1263 ACCAACCGTATACACCGTTTATATGCGGCAAGGACCCGCAAGGAGAGATGCTAGCTAT 1322
Db 1975 ACCAACCGTATACACCGTTTATATGCGGCAAGGACCCGCAAGGAGAGATGCTAGCTAT 2034
QY 1323 CATTTAGCCGGTGGT 1337
Db 2035 CATTTAGCCCTATGAT 2049

```

RESULT 6

```

US-07-703-778D-1
; Sequence 1, Application US/07703778D
; Patent No. 5296366
; GENERAL INFORMATION:
; APPLICANT: Garcia, M.P.E. et al
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
; OF A GENE WHICH CODES FOR STREPTOKINASE, NUCLEOTIDE SEQUENCE
; TITLE OF INVENTION: OBTAINED, RECOMBINANT DNA AND TRANSFORMED MICROORGANISMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Stanger, Michaelson, Spivak and Wallace, Esq.
; STREET: Parkway 109 Office Center, 328 Newman Springs Road,
; CITY: P. O. Box 8489
; CITY: Red Bank
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44Mb IBM compatible diskette
; COMPUTER: IBM PS/2 Model 80
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: Microsoft Word for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/703,778D
; FILING DATE: 19910522
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.
; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: Centro-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)530-6671
; TELEFAX: (908)530-6584
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1245 base pairs
; TYPE: NUCLEOTIDE WITH CORRESPONDING PROTEIN
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis from group C of Lanfield
; ORGANISM: definition

```

```

; IMMEDIATE SOURCE: ATCC-9542 strain
; FEATURE: from 1 to 1245 bp mature peptide
; OTHER INFORMATION: Properties: Streptokinase gene
; OTHER INFORMATION: The gene product binds to human plasminogen
; OTHER INFORMATION: The gene product is an activator of human plasminogen
US-07-703-778D-1
Query Match 67.9%; Score 1127.8; DB 2; Length 1245;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1138; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 183 ATTGCTCGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGGTTCTT 242
Db 1 ATTGCTGGAAGTGGTGGTGTAGACCGTCCATCTGTCAACACAGCAATTTAGTTGTT 60
QY 243 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATATTAGTCTTAAATTTTGAATC 302
Db 61 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATATTAGTCTTAAATTTTGAATC 120
QY 303 GATCTAATCATCAGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
Db 121 GACCTAAATCAACACCTGCTCATGGAGGAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
QY 363 AAACCAATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
Db 181 AAACCAATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 240
QY 423 AAGCTATTCAAGAAACAAATTGATCGCTAAACGTTCCACAGTAAACGACGACTACTTTGAGTTC 482
Db 241 AAGGCTATTCAAGAAACAAATTGATCGCTAAACGTTCCACAGTAAACGACGACTACTTTGAGTTC 300
QY 483 ATTGATTTTGCAGCGATGCAACCAATTTACTGATGCAAAACGCAAGGCTTACTTTGCTGAC 542
Db 301 ATTGATTTTGCAGCGATGCAACCAATTTACTGATGCAAAACGCAAGGCTTACTTTGCTGAC 360
QY 543 AAAGATGTTTCGGTAACTCTGCCACCCCAACCTGTCCAAAGAAATTTTGCCTAAGCGGACAT 602
Db 361 AAAGATGTTTCGGTAACTCTGCCACCCCAACCTGTCCAAAGAAATTTTGCCTAAGCGGACAT 420
QY 603 GTGCGCTTTAGACCATATATAAGAAAAACCAATATACAAAAACCAAGCGAAATCTGTTGATGTG 662
Db 421 GTGCGCTTTAGACCATATATAAGAAAAACCAATATACAAAAACCAAGCGAAATCTGTTGATGTG 480
QY 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCTAAA 722
Db 481 GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCTAAA 540
QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 782
Db 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 600
QY 783 GCTCAAGCACAAGCAATTTTAAAAACCAAAACCAAGCTATAGCATTTATGAAACGTCGAC 842
Db 601 GCTCAAGCACAAGCAATTTTAAAAACCAAAACCAAGCTATAGCATTTATGAAACGTCGAC 660
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGAGTTTTCAGATTTTCAATCGATCAAG 902
Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGAGTTTTCAGATTTTCAATCGATCAAG 720
QY 903 TTTACTTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAATCTGGTCTG 962
Db 721 TTTACTTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 781 AATGAAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATACGTT 1082
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATACGTT 900
QY 1083 GATGTCGATACCAACGAATTTGCTTAAAAAGTGACGCTCTTAAACGCTAGCGAACGTAAC 1142

```

Db 901 GATGTCAACACCAAGCAATTTGCTTAAAGCGAGCAGCTCTTAAAGCTAGCGAAGCTAAC 960
QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGTGATAAGGCTAACTACTCTACACAACTCTC 1202
Db 961 TTAGACTTCAGAGATTATACGATCCTCGTGTGATAAGGCTAACTACTCTACACAACTCTC 1020
QY 1203 GATGCTTTTGGTATTATGGAATCTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGGAATCTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1080
QY 1263 ACCAAGCGTATCATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1322
Db 1081 ACCAAGCGTATCATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCGTATGAT 1155

RESULT 7

US-07-854-596B-25
; Sequence 25, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1257
; OTHER INFORMATION: /note= "Methionyl-streptokinase"
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1248
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 4..1248
US-07-854-596B-25

Query Match 67.9%; Score 1127.6; DB 2; Length 1257;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 180 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCAATCTGTCAACACAGCCAAATTTAGTT 239
Db 4 ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCAATCTGTCAACACAGCCAAATTTAGTT 63
QY 240 GTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGA 299
Db 64 GTTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGA 123
QY 300 ATCGATCTAAACATCAGACCTCTCATGCGAGAAAGACAGACAGGCTTAAAGTCCAAA 359
Db 124 ATTGACCTTAAACATCAGACCTCTCATGCGAGAAAGACAGACAGGCTTAAAGTCCAAA 183
QY 360 TCAAAACCATTTGCTACTGATAGTGGCGATGTCAATATAAATTTGAGAAAGCTGACTTA 419
Db 184 TCAAAACCATTTGCTACTGATAGTGGCGATGTCAATATAAATTTGAGAAAGCTGACTTA 243
QY 420 CTAAAGGCTATTCAAGAAACAATTGATTCGCTAACGTCACAGTAACGACGACTCTTTGAG 479
Db 244 CTAAAGGCTATTCAAGAAACAATTGATTCGCTAACGTCACAGTAACGACGACTCTTTGAG 303
QY 480 GTCATTGATTTTGCAGGCGATGCAACCAATTAATCTGATCGAAACGCAAGGCTCTACTTTGCT 539
Db 304 GTCATTGATTTTGCAGGCGATGCAACCAATTAATCTGATCGAAACGCAAGGCTCTACTTTGCT 363
QY 540 GACAAAGATGTTTCGTTAACTTTCGCGACCCCACTGCTCCCAAGAAATTTTGTAAAGCGA 599
Db 364 GACAAAGATGTTTCGTTAACTTTCGCGACCCCACTGCTCCCAAGAAATTTTGTAAAGCGA 423
QY 600 CATGTGCGGCTTAGACCATATAAAGAAAAACAATACAAAATCAAGCGAAATCTGTTGAT 659
Db 424 CATGTGCGGCTTAGACCATATAAAGAAAAACAATACAAAATCAAGCGAAATCTGTTGAT 483
QY 660 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACGAGTCTC 719
Db 484 GTGGAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACGAGTCTC 543
QY 720 AAAGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 779
Db 544 AAAGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 603
QY 780 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGT 839
Db 604 CTAGCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGT 663
QY 840 GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAA 899
Db 664 GACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAA 723
QY 900 GAGTTTACTTACCGTGTATAAAATCGGAAACAGCTTATAGATCAATATAAAATCTCGT 959
Db 724 GAGTTTACTTACCGTGTATAAAATCGGAAACAGCTTATAGATCAATATAAAATCTCGT 783
QY 960 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAA 1019
Db 784 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAA 843
QY 1020 GGGGAAAGCGGTATGATCCCTTTGATCGAGTCACTTTGAAATCTGTTCCACCATCAATAC 1079
Db 844 GGGGAAAGCGGTATGATCCCTTTGATCGAGTCACTTTGAAATCTGTTCCACCATCAATAC 903
QY 1080 GTTGATGTCGATACCAACGATTTGCTTAAAAAGTGGAGCTCTTAAAGCTAGCGAAAGT 1139
Db 904 GTTGATGTCGATACCAACGATTTGCTTAAAAAGTGGAGCTCTTAAAGCTAGCGAAAGT 963
QY 1140 AACTTTAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAAACTACTCTTACAACT 1199
Db 964 AACTTTAGACTTCAGAGATTTTATACGATCCTCGTATAGGCTTAAACTACTCTTACAACT 1023

QY 1200 CTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATAATCACGAT 1259
 Db 1024 CTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATAATCACGAT 1083
 QY 1260 GACACCAACCGTATCATTAACCGTTTATATGGGCAAGCGACCGAGAGAGAGTGTAGC 1319
 Db 1084 GACACCAACCGTATCATTAACCGTTTATATGGGCAAGCGACCGAGAGAGAGTGTAGC 1143
 QY 1320 TATCAATTTAGCGGTGT 1337
 Db 1144 TATCAATTTAGCCTATGAT 1161

RESULT 8

US-07-854-596B-27
 ; Sequence 27, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 27:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1512 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..1512
 ; OTHER INFORMATION: /note= "Streptokinase
 ; OTHER INFORMATION: fused to a yeast alpha-factor"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 7..1503
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 7..1503
 ; US-07-854-596B-27

Query Match 67.8%; Score 1126.2; DB 2; Length 1512;
 Best Local Similarity 98.0%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 175 ATACCATGATTTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAT 234

Db 254 ATAAAAAGAAATTCGTGGACCTCGAGTGGCTTAGACCGTCCATCTGTCTCAACAACAGCCAAT 313
 QY 235 TGGTTGTTAGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 294
 Db 314 TAGTTGTTAGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTT 373
 QY 295 TTGAAATCGATCTAACATCAAGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAAGTC 354
 Db 374 TTGAAATGACCTAACATCAAGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAAGTC 433
 QY 355 CAAATCAAAACCAATTTGCTTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTG 414
 Db 434 CAAATCAAAACCAATTTGCTTACTGATAGTGGCGGATGCCACATAAACTTGAGAAAGCTG 493
 QY 415 ACTTACTAAAGGCTATTCAAGAACAAATTCGCTTAAGCTCCACAGTAACACGACGACTACT 474
 Db 494 ACTTACTAAAGGCTATTCAAGAACAAATTCGCTTAAGCTCCACAGTAACACGACGACTACT 553
 QY 475 TTGAGGTCAATTGATTTTGCAGCGATGCAACCATTAATCTGATCGAAACCGCAAGGTCTACT 534
 Db 554 TTGAGGTCAATTGATTTTGCAGCGATGCAACCATTAATCTGATCGAAACCGCAAGGTCTACT 613
 QY 535 TTGCTGACAAAGATGGTTGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTCTAA 594
 Db 614 TTGCTGACAAAGATGGTTGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTCTAA 673
 QY 595 GCGGACATGTGCGGTTAGACCATATAAGAAACCAATACAAACCAACCAAGCGAAATCTG 654
 Db 674 GCGGACATGTGCGGTTAGACCATATAAGAAACCAATACAAACCAACCAAGCGAAATCTG 733
 QY 655 TTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCCAG 714
 Db 734 TTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCCAG 793
 QY 715 GTCTCAAAGATACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAG 774
 Db 794 GTCTCAAAGATACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAG 853
 QY 775 AATTACTAGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATG 834
 Db 854 AATTACTAGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATG 913
 QY 835 AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGG 894
 Db 914 AACGTGACTCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGG 973
 QY 895 ATCAAGAGTTTACTTTACCGTGTAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAT 954
 Db 974 ATCAAGAGTTTACTTTACCATGTCAAAATCGGGAACAAAGCTTTATGAGATCAATAAAAAT 1033
 QY 955 CTGCTGTAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATTACGTCCTTA 1014
 Db 1034 CTGCTGTAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATTACGTCCTTA 1093
 QY 1015 AAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACCTGAAACCTGTTCCACATCA 1074
 Db 1094 AAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACCTGAAACCTGTTCCACATCA 1153
 QY 1075 AATACGTTGATGTGATACCAACGAATTTGCTTAAAGAGTGAGCAGCTCTTAAACAGTAGCG 1134
 Db 1154 AATACGTTGATGTCAACACCAACGAATTTGCTTAAAGAGTGAGCAGCTCTTAAACAGTAGCG 1213
 QY 1135 AACGTAACTTAGACTTCAGAGATTTATACGATCCCTGCTGTGATAAGGCTTAACTCTCTACA 1194
 Db 1214 AACGTAACTTAGACTTCAGAGATTTATACGATCCCTGCTGTGATAAGGCTTAACTCTCTACA 1273
 QY 1195 ACAATCTCGATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGGATAATC 1254
 Db 1274 ACAATCTCGATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGGATAATC 1333
 QY 1255 ACAGTACACCAACCGCTATCATAACCGTTTATATGGGCAAGCGACCCCAAGGAGAGAATG 1314
 Db 1334 ACAGTACACCAACCGCTATCATAACCGTTTATATGGGCAAGCGACCCCAAGGAGAGAATG 1393

QY 1315 CTAGCTATCATTTAGCCGGTGGT 1337
Db 1394 CTAGCTATCATTTAGCCCTATGAT 1416

RESULT 9

US-07-854-596B-18
; Sequence 18, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1317
; OTHER INFORMATION: /note= "OmpA fused to mature
; OTHER INFORMATION: streptokinase gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1308
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 4..1308
US-07-854-596B-18

Query Match 67.7%; Score 1124.6; DB 2; Length 1317;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCCCAATTGGTTGTT 242
Db 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCCCAATTAGTTGTT 126
QY 243 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 127 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 186
QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACGAGCGCTTAAGTCCAAATCA 362

Db 187 GACCTAACATCAGCACTGCTCATGGAGGAAAGACAGACGAGCGCTTAAGTCCAAATCA 246
QY 363 AAACCATTTGCTTACTGATAGTGGCGGATGTACATATAACTTTGAGAAAGCTGACTACTA 422
Db 247 AAACCATTTGCTTACTGATAGTGGCGGATGTACATATAACTTTGAGAAAGCTGACTACTA 306
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGAGCACTACTTTTCAGGTC 482
Db 307 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGAGCACTACTTTTCAGGTC 366
QY 483 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGCGCAAGGTCTACTTTCTGAC 542
Db 367 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGCGCAAGGTCTACTTTCTGAC 426
QY 543 AAAGATGCTTGGTAAACCTTCGCCGACCCAACTGTCGCAAGATTTTTCGTAAGCGGACAT 602
Db 427 AAAGATGCTTGGTAAACCTTCGCCGACCCAACTGTCGCAAGATTTTTCGTAAGCGGACAT 486
QY 603 GTGGCGGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 662
Db 487 GTGGCGGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 546
QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGGTCTCAA 722
Db 547 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCGAGGTCTCAA 606
QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACGATCATCTCAAGAAATTACTA 782
Db 607 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACGATCATCTCAAGAAATTACTA 666
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 667 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 726
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 902
Db 727 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 786
QY 903 TTTTACTTACCGTTTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 787 TTTTACTTACCGTTTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 846
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 1022
Db 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 906
QY 1023 GAAAAGCGGTATGATCCCTTTTATGTCGAGTCACTTTGAAACTGTTCCACCATCAAAATACGTT 1082
Db 907 GAAAAGCGGTATGATCCCTTTTATGTCGAGTCACTTTGAAACTGTTCCACCATCAAAATACGTT 966
QY 1083 GATGTCGATACCAACGAATTTGCTTAAAGAGTGAAGAGTCTTTAAACAGCTAGGAAACGTAAC 1142
Db 967 GATGTCGATACCAACGAATTTGCTTAAAGAGTGAAGAGTCTTTAAACAGCTAGGAAACGTAAC 1026
QY 1143 TTAGACTTCAGAGATTTTATACGATCTCTGATAGGCTAAACTTACTCTCAACAATCTC 1202
Db 1027 TTAGACTTCAGAGATTTTATACGATCTCTGATAGGCTAAACTTACTCTCAACAATCTC 1086
QY 1203 GATCTCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGGAGTAATCAGGATGAC 1262
Db 1087 GATCTCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGGAGTAATCAGGATGAC 1146
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGGAAGGAGAGAAATGCTAGCTAT 1322
Db 1147 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGGAAGGAGAGAAATGCTAGCTAT 1206
QY 1323 CATTTAGCCGGTGGT 1337
Db 1207 CATTTAGCCCTATGAT 1221

RESULT 10

US-07-854-596B-14
; Sequence 14, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1335 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1335
; OTHER INFORMATION: /note= "Streptokinase gene from S."
; OTHER INFORMATION: equisimilis"
; NAME/KEY: CDS
; LOCATION: 7..1326
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 7..1326
; US-07-854-596B-14

Query Match 67.7%; Score 1124.6; DB 2; Length 1335;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 242
Db 85 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 144
QY 243 AGCGTGTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 145 AGCGTGTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 204
QY 303 GATCTAAATCAGACCTGCTCATGGAGAAAGACAGACGCTTAACTCCAAATCA 362
Db 205 GACCTAAATCAGACCTGCTCATGGAGAAAGACAGACGCTTAACTCCAAATCA 264
QY 363 AAACCAATTTGCTAGTGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 265 AAACCAATTTGCTAGTGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 324

QY 423 AAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTTC 482
Db 325 AAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTTC 384
QY 483 ATTGATTTTGAACGGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
Db 385 ATTGATTTTGAACGGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 444
QY 543 AAAGATGTTTGGTAACTTCCGCGACCAACCTGTCGAAGAAATTTTTCGTAAGCGGACAT 602
Db 445 AAAGATGTTTGGTAACTTCCGCGACCAACCTGTCGAAGAAATTTTTCGTAAGCGGACAT 504
QY 603 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 505 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 564
QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACCGAGTCTCAA 722
Db 565 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTTCAGACCGAGTCTCAA 624
QY 723 GATATTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 782
Db 625 GATATTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 684
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCGCTATACGATTTATGAACGTGAC 842
Db 685 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCGCTATACGATTTATGAACGTGAC 744
QY 843 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 745 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 804
QY 903 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAATCTGGTCTG 962
Db 805 TTTACTTTACCATGTCAAAAAATCGGGAAACAGCTTATAGGATCAATAAAAATCTGGTCTG 864
QY 963 AATGAAGAAATAAACACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 865 AATGAAGAAATAAACACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 924
QY 1023 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACCTTGAAACCTGTTTCAACCATCAATACGTT 1082
Db 925 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACCTTGAAACCTGTTTCAACCATCAATACGTT 984
QY 1083 GATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC 1142
Db 985 GATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC 1044
QY 1143 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAATAGGCTAAACTCTTACCAATCTC 1202
Db 1045 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAATAGGCTAAACTCTTACCAATCTC 1104
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1262
Db 1105 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGCATGAC 1164
QY 1263 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1322
Db 1165 ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAAATGCTAGCTAT 1224
QY 1323 CATTTAGCCGGTGT 1337
Db 1225 CATTTAGCCCTATGAT 1239

RESULT 11
US-07-854-596B-42
; Sequence 42, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1458 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..1458
 ; OTHER INFORMATION: /note= "Hirudin-streptokinase
 ; OTHER INFORMATION: fusion linked by Factor Xa cleavable IEGR"
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1449
 ; NAME/KEY: mat_peptide
 ; LOCATION: 1..1449
 ; US-07-854-596B-42

Query Match 67.7%; Score 1124.6; DB 2; Length 1458;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY	183	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCCAATTGGTTGTT	242
Db	208	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCCAATTGGTTGTT	267
QY	243	AGCGTTGCTGCTACTGTGTGGGGGCAACATCAAGACATTAGTCTTAATTTTGAATC	302
Db	268	AGCGTTGCTGCTACTGTGTGGGGGCAACATCAAGACATTAGTCTTAATTTTGAATC	327
QY	303	GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCGTCTTAAGTCCAAATCA	362
Db	328	GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCGTCTTAAGTCCAAATCA	387
QY	363	AAACCAATTGCTACTGATAGTGGCGGATGTCAATAACTTGAAGAAGCTGACTTACTA	422
Db	388	AAACCAATTGCTACTGATAGTGGCGGATGTCAATAACTTGAAGAAGCTGACTTACTA	447
QY	423	AAGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTACGACGACTTGTGAGTC	482
Db	448	AAGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTACGACGACTTGTGAGTC	507
QY	483	ATTGATTTTGAACCGATGCAACATTACTGATCGAAACGCAAGGCTCTCTTTGCTGAC	542
Db	508	ATTGATTTTGAACCGATGCAACATTACTGATCGAAACGCAAGGCTCTCTTTGCTGAC	567

QY	543	AAAGATGGTTTCGGTAAACCTTGGCGACCCCAACCTGTCCCAAGAAATTTTGTGAAGCGGACAT	602
Db	568	AAAGATGGTTTCGGTAAACCTTGGCGACCCCAACCTGTCCCAAGAAATTTTGTGAAGCGGACAT	627
QY	603	GTGCGGTTAGACCATATAAAGAAAAACAATACAAAACCAAGCGGAAATCTCTTGTATGTG	662
Db	628	GTGCGGTTAGACCATATAAAGAAAAACAATACAAAACCAAGCGGAAATCTCTTGTATGTG	687
QY	663	GAATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGGATTTACAGACCGGTCTCAAA	722
Db	688	GAATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGGATTTACAGACCGGTCTCTCAA	747
QY	723	GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCACAATCTCAAGAAATTA	782
Db	748	GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCACAATCTCAAGAAATTA	807
QY	783	GCTCAAGCACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAAACGTAC	842
Db	808	GCTCAAGCACAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAAACGTAC	867
QY	843	TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	902
Db	868	TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	927
QY	903	TTTACTTACCGGTAAATAATCGGCAACAAGCTTATAGGATCAATAAATAATCTCGTCTG	962
Db	928	TTTACTTACCGGTAAATAATCGGCAACAAGCTTATAGGATCAATAAATAATCTCGTCTG	987
QY	963	AATGAAGAAATAAACAACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG	1022
Db	988	AATGAAGAAATAAACAACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG	1047
QY	1023	GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACATCAATACGTT	1082
Db	1048	GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACATCAATACGTT	1107
QY	1083	GATGTCGATACCAAGAAATGCTTAAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAAC	1142
Db	1108	GATGTCGATACCAAGAAATGCTTAAAGAGTGAAGCTCTTAAACAGCTAGCGAAGCTAAC	1167
QY	1143	TTAGACTTTCAGAGATTTATACGATCCTCGTATAGGCTTAAACTACTCTACAAACATCTC	1202
Db	1168	TTAGACTTTCAGAGATTTATACGATCCTCGTATAGGCTTAAACTACTCTACAAACATCTC	1227
QY	1203	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATAATCACGATGAC	1262
Db	1228	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATAATCACGATGAC	1287
QY	1263	ACCAACCGTATCATTAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGAAATGCTAGCTAT	1322
Db	1288	ACCAACCGTATCATTAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGAAATGCTAGCTAT	1347
QY	1323	CATTAGCCGGTGGT 1337	
Db	1348	CATTAGCCGTATGAT 1362	

RESULT 12
 US-07-854-596B-46
 ; Sequence 46, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL

```

; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1467 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1467
; OTHER INFORMATION: /note= "Streptokinase-hirudin
; OTHER INFORMATION: fusion linked by Factor Xa-cleavable IEGR"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1449
; NAME/KEY: mat_peptide
; LOCATION: 1..1449
; US-07-854-596B-46

Query Match 67.7%; Score 1124.6; DB 2; Length 1467;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACGGTCCATCTGTCAACACGCAATGGTGT 242
DB 1 ATTGCTGGACCTGAGTGGCTGCTAGACGGTCCATCTGTCAACACGCAATGGTGT 60

QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
DB 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 120

QY 303 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA 362
DB 121 GACCTAACATCACGACCTGCTCATGAGGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA 180

QY 363 AAACCAATTTCTACTGATGAGGCGCATGTCAATAACTTGGAAGCTGACTTACTA 422
DB 181 AAACCAATTTCTACTGATGAGGCGCATGTCAATAACTTGGAAGCTGACTTACTA 240

QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAAGCAAGCTACTTTGAGTC 482
DB 241 AAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAAGCAAGCTACTTTGAGTC 300

QY 483 ATTGATTTTGAACGCGATGCAACCAATTTACTGATCGAAACGCAAGGCTACTTTGCTGAC 542
DB 301 ATTGATTTTGAACGCGATGCAACCAATTTACTGATCGAAACGCAAGGCTACTTTGCTGAC 360

QY 543 AAAGATGGTTCGGTAACTTCCGACCACTGTGCAAGAAATTTTGTCTAAGCGGACAT 602
DB 361 AAAGATGGTTCGGTAACTTCCGACCACTGTGCAAGAAATTTTGTCTAAGCGGACAT 420

QY 603 GTSCGGTGTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662

```

```

DB 421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
QY 663 GAATATATCTGTACAGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCTCAA 722
DB 481 GAATATATCTGTACAGATTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCTCAA 540
QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 782
DB 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 600
QY 783 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
DB 601 GCTCAAGCACAAGCAATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 843 TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAG 902
DB 661 TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAG 720
QY 903 TTTACTTTACCGTGTAAATAATCGGGAACAACCTTATAGGATCAATAAAAAATCTGGTCTG 962
DB 721 TTTACTTTACCATGTCAAAATATCGGGAACAACCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATPAAACACACCTGATCTCTGAGAAATATTAGTCTTTAAAAAAGG 1022
DB 781 AATGAAGAAATPAAACACACCTGATCTCTGAGAAATATTAGTCTTTAAAAAAGG 840
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAACCTGTTTCAACCATCAATACGTT 1082
DB 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAACCTGTTTCAACCATCAATACGTT 900
QY 1083 GATGTCGATACCAACGAATTCCTAAAGAGTCAAGAGCTCTTTAAACAGCTAGCGCAACGTAAAC 1142
DB 901 GATGTCGATACCAACGAATTCCTAAAGAGTCAAGAGCTCTTTAAACAGCTAGCGCAACGTAAAC 960
QY 1143 TTAGACTTTCAGAGATTTTATACGATCTCTGATGATGAGCTAAACTACTCTACACCAATCTC 1202
DB 961 TTAGACTTTCAGAGATTTTATACGATCTCTGATGATGAGCTAAACTACTCTACACCAATCTC 1020
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCACGATGAC 1262
DB 1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCACGATGAC 1080
QY 1263 ACCAACGCTATCATAAACCGTTTATATGGCAAGCGACCCGAGGAGAGAGAAATGCTAGCTAT 1322
DB 1081 ACCAACGCTATCATAAACCGTTTATATGGCAAGCGACCCGAGGAGAGAGAAATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGT 1337
DB 1141 CATTTAGCCGTATGAT 1155

```

```

RESULT 13
US-07-854-596B-34
; Sequence 34, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 2589 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2589
OTHER INFORMATION: /note=
OTHER INFORMATION: "OmpA-Streptokinase-streptokinase fusion linked by thrombin-cleavable VELQVVRG"
FEATURE:
NAME/KEY: CDS
LOCATION: 4..2580
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 4..2580
US-07-854-596B-34
Query Match i 67.7%; Score 1124.6; DB 2; Length 2589;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 242
DB 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 126
QY 243 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
DB 127 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAAT 186
QY 303 GATCTAATCATCAGCACTGCTCATGAGGAGAAAGACAGACAGGCTTAAAGTCCAAATCA 362
DB 187 GACCTAATCATCAGCACTGCTCATGAGGAGAAAGACAGACAGGCTTAAAGTCCAAATCA 246
QY 363 AAAACATTGCTACTGATAGTGGCGGATGCTCATATAAATTTGAGAAAGCTGACTTACTA 422
DB 247 AAAACATTGCTACTGATAGTGGCGGATGCTCATATAAATTTGAGAAAGCTGACTTACTA 306
QY 423 AAGGCTATTCAAGAACAAATGATCGCTTAACGTCACAGTACACGACGACTACTTTGAGGTC 482
DB 307 AAGGCTATTCAAGAACAAATGATCGCTTAACGTCACAGTACACGACGACTACTTTGAGGTC 366
QY 483 ATTGATTTTGAAGCGATGCAACCATTAATGATCGAAACGGCAAGGCTTACTTTGCTGAC 542
DB 367 ATTGATTTTGAAGCGATGCAACCATTAATGATCGAAACGGCAAGGCTTACTTTGCTGAC 426
QY 543 AAGATGTTTCGGTACCTTTCGCGCCACCTGCTCAAGAAATTTTGTCTAAGCGGACAT 602
DB 427 AAGATGTTTCGGTACCTTTCGCGCCACCTGCTCAAGAAATTTTGTCTAAGCGGACAT 486
QY 603 GTGCGGTTAGACATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
DB 487 GTGCGGTTAGACATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 546
QY 663 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 722
DB 547 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 606

QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACAATCTCAAGAAATTACTA 782
DB 607 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACAATCTCAAGAAATTACTA 666
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAAAGTGC 842
DB 667 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAAAGTGC 726
QY 843 TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
DB 727 TCCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 786
QY 903 TTTACTTACCGTGTATAAAATCGGAAACAAAGCTTATAGGATCAATATAAAATCTCGTCTG 962
DB 787 TTTACTTACCATGTCAAAATCGGAAACAAAGCTTATGAGATCAAATAAAATCTCGTCTG 846
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG 1022
DB 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGG 906
QY 1023 GAAAAGCGGTATGATCCCTTTTGTCGCACTCACTTTGAAACTGTTCCACCATCAAAATACGTT 1082
DB 907 GAAAAGCGGTATGATCCCTTTTGTCGCACTCACTTTGAAACTGTTCCACCATCAAAATACGTT 966
QY 1083 GATGTCGATACCAACGAATTCGTAATAAGTGAAGAGCTCTTAAACAGCTAGCGAAACGTAAC 1142
DB 967 GATGTCGATACCAACGAATTCGTAATAAGTGAAGAGCTCTTAAACAGCTAGCGAAACGTAAC 1026
QY 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTGAATAGGCTTAACTCTCTACAAATCTC 1202
DB 1027 TTAGACTTCAGAGATTTTATACGATCCTCGTGAATAGGCTTAACTCTCTACAAATCTC 1086
QY 1203 GATGCTTTTGGTATTATGCACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1262
DB 1087 GATGCTTTTGGTATTATGCACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1146
QY 1263 ACCAACCGGTATCATAACCGTTTATATGGGCAAGGACCCGAGGAGAGAGATGCTAGCTAT 1322
DB 1147 ACCAACCGGTATCATAACCGTTTATATGGGCAAGGACCCGAGGAGAGAGATGCTAGCTAT 1206
QY 1323 CATTTAGCGGTGGT 1337
DB 1207 CATTTAGCGGTATGAT 1221
RESULT 14
US-09-374-038-13
Sequence 13, Application US/09374038
Patent No. 6309873
GENERAL INFORMATION:
APPLICANT: Madrazo, Isis Del Carmen Torrens
APPLICANT: Garcia, Jose De Jesus De La Fuente
APPLICANT: Ojalvo, Ariana Garcia
APPLICANT: Menendez, Alina Seraleña
APPLICANT: Escalona, Elder Pupo
APPLICANT: Masso, Julio Raul Fernandez
APPLICANT: Griego, Martha De Jesus Gonzalez
TITLE OF INVENTION: STREPTOKINASE MUTANTS
FILE REFERENCE: Sequence Listings 1-14 re: 976-5
Patent No. 6309873
CURRENT APPLICATION NUMBER: US/09/374,038
CURRENT FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 1122
TYPE: DNA
ORGANISM: Streptococcus equisimilis
US-09-374-038-13
Query Match 66.0%; Score 1096.6; DB 3; Length 1122;
Best Local Similarity 98.7%; Pred. No. 5e-308;
Matches 1105; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 180 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTGGTT 239
 Db 1 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 60
 Qy 240 GTTAGCGTGGTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 299
 Db 61 GTTAGCGTGGTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 120
 Qy 300 ATCGATCTAACATCAGACCTGCTATGAGGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 359
 Db 121 ATGACCTTAACATCAGACCTGCTATGAGGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 180
 Qy 360 TCAAAACCAATTTGCTACTGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 419
 Db 181 TCAAAACCAATTTGCTACTGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 240
 Qy 420 CTAAAGGCTATTCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAG 479
 Db 241 CTAAAGGCTATTCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAG 300
 Qy 480 GTCAATGATTTTGAAGCAATGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 539
 Db 301 GTCAATGATTTTGAAGCAATGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 360
 Qy 540 GACAAAGATGGTTCGCTAACCTTGGCGACCTGTCACCAATTTTGGCTTAAAGCGGA 599
 Db 361 GACAAAGATGGTTCGCTAACCTTGGCGACCTGTCACCAATTTTGGCTTAAAGCGGA 420
 Qy 600 CATGTGCGGTTAGACCATATAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 659
 Db 421 CATGTGCGGTTAGACCATATAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 480
 Qy 660 GTGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGGCTC 719
 Db 481 GTGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGGCTC 540
 Qy 720 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCAATCTCAAGAAATTA 779
 Db 541 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCAATCTCAAGAAATTA 600
 Qy 780 CTAGCTCAAGCAACAAAGCAATTTAAACAAACCAAGCAAGCTTATGACAGT 839
 Db 601 CTAGCTCAAGCAACAAAGCAATTTAAACAAACCAAGCAAGCTTATGACAGT 660
 Qy 840 GACTCTCAATCTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 899
 Db 661 GACTCTCAATCTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 720
 Qy 900 GAGTTTACTTACCGTGTAAATCGGGAACAGCTTATAGGATCAATAAATCTGGT 959
 Db 721 GAGTTTACTTACCGTGTAAATCGGGAACAGCTTATAGGATCAATAAATCTGGT 780
 Qy 960 CTGAATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGCTTAAATAA 1019
 Db 781 CTGAATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGCTTAAATAA 840
 Qy 1020 GGGGAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAATATC 1079
 Db 841 GGGGAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAATATC 900
 Qy 1080 GTTGATGTCGATCCCAAGAAATGCTTAAAGTGACAGCTCTTAAAGTGACAGTACGAGT 1139
 Db 901 GTTGATGTCGATCCCAAGAAATGCTTAAAGTGACAGCTCTTAAAGTGACAGTACGAGT 960
 Qy 1140 AACTTAGACTTCAGAGATTTATAGGATCTCTGAGTAAAGCTAAATCTACTCTCAACAAT 1199
 Db 961 AACTTAGACTTCAGAGATTTATAGGATCTCTGAGTAAAGCTAAATCTACTCTCAACAAT 1020
 Qy 1200 CTGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTATAGGAGTAAATCAGAT 1259
 Db 1021 CTGATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTATAGGAGTAAATCAGAT 1080

Qy 1260 GACACCAACCGTATCATAACCGTTTATATGGCAAGCGA 1298
 Db 1081 GACACCAACCGTATCATAACCGTTTATATGGCAAGCGA 1119
 RESULT 15
 US-09-658-179-13
 ; Sequence 13, Application US/09658179
 ; Patent No. 6413759
 ; GENERAL INFORMATION:
 ; APPLICANT: Madrazo, Isis Del Carmen Torrens
 ; APPLICANT: Garcia, Jose De Jesus De La Fuente
 ; APPLICANT: Ojalvo, Ariana Garcia
 ; APPLICANT: Menendez, Alina Seralena
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6413759
 ; CURRENT APPLICATION NUMBER: US/09/658,179
 ; CURRENT FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 1122
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; US-09-658-179-13

Query Match 66.0%; Score 1096.6; DB 3; Length 1122;
 Best Local Similarity 98.7%; Pred. No. 5e-308;
 Matches 1105; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 Qy 180 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTGGTT 239
 Db 1 ATGATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTT 60
 Qy 240 GTTAGCGTGGTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 299
 Db 61 GTTAGCGTGGTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAA 120
 Qy 300 ATCGATCTAACATCAGACCTGCTATGAGGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 359
 Db 121 ATGACCTTAACATCAGACCTGCTATGAGGGAAGACAGACGAGCAAGCTTAAAGTCCAAA 180
 Qy 360 TCAAAACCAATTTGCTACTGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 419
 Db 181 TCAAAACCAATTTGCTACTGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 240
 Qy 420 CTAAAGGCTATTCAAGAACAAATGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 479
 Db 241 CTAAAGGCTATTCAAGAACAAATGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 300
 Qy 480 GTCAATGATTTTGAAGCAATGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 539
 Db 301 GTCAATGATTTTGAAGCAATGATGAGTGGCGGATGTCATATAAATCTGAGAAAGCTGACTTA 360
 Qy 540 GACAAAGATGGTTCGCTAACCTTGGCGACCTGTCACCAATTTTGGCTTAAAGCGGA 599
 Db 361 GACAAAGATGGTTCGCTAACCTTGGCGACCTGTCACCAATTTTGGCTTAAAGCGGA 420
 Qy 600 CATGTGCGGTTAGACCATATAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 659
 Db 421 CATGTGCGGTTAGACCATATAAGAAACCAATAACAAACCAAGCGAAATCTGTTGAT 480
 Qy 660 GTGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGGCTC 719
 Db 481 GTGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGGCTC 540
 Qy 720 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCAATCTCAAGAAATTA 779
 Db 541 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACCAATCTCAAGAAATTA 600

```
QY 780 CTAGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGT 839
Db |||||||
QY 601 CTAGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGT 660
Db |||||||
QY 840 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 899
Db |||||||
QY 661 GACTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAA 720
Db |||||||
QY 900 GAGTTTACTTACCGTGTAAATAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGT 959
Db |||||||
QY 721 GAGTTTACTTACCATGTCAAAATTCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGT 780
Db |||||||
QY 960 CTGAATGAAGAAATAAAACAAACACCTGACCTGATCTCTGAGAAATATTAGTCTTTAAAAAA 1019
Db |||||||
QY 781 CTGAATGAAGAAATAAAACAAACACCTGACCTGATCTCTGAGAAATATTAGTCTTTAAAAAA 840
Db |||||||
QY 1020 GGGGAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAAACTGTTCAACCATCAATAC 1079
Db |||||||
QY 841 GGGGAAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAAACTGTTCAACCATCAATAC 900
Db |||||||
QY 1080 GTTGATGTGATACCAAGAAATGCTAAAGTGAGCAGCTCTTAAACAGCTAGCGACGT 1139
Db |||||||
QY 901 GTTGATGTCAACCAACGAATGCTAAAGCGAGCAGCTCTTAAACAGCTAGCGACGT 960
Db |||||||
QY 1140 AACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAAT 1199
Db |||||||
QY 961 AACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAAT 1020
Db |||||||
QY 1200 CTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGAT 1259
Db |||||||
QY 1021 CTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGAT 1080
Db |||||||
QY 1260 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGA 1298
Db |||||||
QY 1081 GACACCAACCGTATCATACCGTTTATATGGGCAAGCGA 1119
Db |||||||
```

Search completed: February 1, 2006, 12:43:42
Job time : 294.256 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 17:06:30 ; Search time 1372.44 Seconds
(without alignments)
10008.068 Million cell updates/sec

Title: US-09-940-235-10

Perfect score: 1661

Sequence: 1 gcaaccccgccagcttagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	100.0	1661	3	US-09-940-235-10
2	1282.2	77.2	1377	3	US-09-940-235-5
3	1219.4	73.4	1327	3	US-09-940-235-6
4	1185	71.3	1541	3	US-09-940-235-9
5	1165.8	70.2	2096	3	US-09-940-235-12
6	1150.2	69.2	1782	3	US-09-940-235-11
7	1148.6	69.2	1245	3	US-09-940-235-1
8	995	59.9	1323	8	US-10-474-792-657
9	316.8	19.1	6988	7	US-10-236-392-1
10	316.8	19.1	7679	9	US-10-831-704-38
11	316.8	19.1	7680	3	US-09-964-824A-574
12	316.8	19.1	7680	5	US-10-171-311-63
13	316.8	19.1	7680	6	US-10-236-031B-69
14	316.8	19.1	7680	6	US-10-374-979-75
15	316.8	19.1	7680	7	US-10-182-936A-75
16	316.8	19.1	7680	7	US-10-641-643-1289
17	316.8	19.1	7680	7	US-10-717-597-222
18	316.8	19.1	7680	8	US-10-788-792-79
19	316.8	19.1	7680	8	US-10-477-238A-654
20	316.8	19.1	7680	8	US-10-680-287A-654
21	316.8	19.1	7680	8	US-10-278-698-88
22	316.8	19.1	7680	8	US-10-278-698-603
23	316.8	19.1	7680	9	US-10-843-641A-5877
					Sequence 10, Appl
					Sequence 5, Appl
					Sequence 6, Appl
					Sequence 9, Appl
					Sequence 12, Appl
					Sequence 11, Appl
					Sequence 1, Appl
					Sequence 657, App
					Sequence 1, Appl
					Sequence 38, Appl
					Sequence 574, App
					Sequence 63, Appl
					Sequence 69, Appl
					Sequence 75, Appl
					Sequence 75, Appl
					Sequence 1289, App
					Sequence 222, App
					Sequence 79, Appl
					Sequence 654, App
					Sequence 654, App
					Sequence 88, Appl
					Sequence 603, App
					Sequence 5877, Ap

ALIGNMENTS

RESULT 1

US-09-940-235-10

; Sequence 10, Application US/09940235

; Publication No. US20030059921A1

; GENERAL INFORMATION:

; APPLICANT: Kumar, Rajesh

; APPLICANT: Sahni, Girish

; APPLICANT: Roy, Chait

; APPLICANT: Rajagopal, Kammara

; APPLICANT: Nihalani, Deepak

; APPLICANT: Sundaram, Vasudha

; APPLICANT: Yadav, Mahavir

; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE

; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION

; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID

; TITLE OF INVENTION: PROTEIN

; FILE REFERENCE: 07064-009002

; CURRENT APPLICATION NUMBER: US/09/940,235

; CURRENT FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 09/471,349

; PRIOR FILING DATE: 1999-12-23

; PRIOR APPLICATION NUMBER: IN 3825/DEL/98

; PRIOR FILING DATE: 1998-12-24

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 1661

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Hybrid cassette

US-09-940-235-10

Query Match 100.0%; Score 1661; DB 3; Length 1661;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACCCCGCCAGCCTAGCCGGGTCCTCAACGACGAGCAGCATGCGCACCCCGTGG 60

Db 1 GCACCCCGCCAGCCTAGCCGGGTCCTCAACGACGAGCAGCATGCGCACCCCGTGG 60

Qy 61 CCAGGACCCCAACCGTCCCGAGATCTCGATCCCGGAAATTAATACGACTCATTATAGG 120

Db 61 CCAGGACCCCAACCGTCCCGAGATCTCGATCCCGGAAATTAATACGACTCATTATAGG 120

Qy 121 AGACCACACCGGTTTCCTCTAGAAATAATTTTGTAAAGAGGATATACCA 180

Db 121 AGACCACACCGGTTTCCTCTAGAAATAATTTTGTAAAGAGGATATACCA 180

Qy 121 AGACCACACCGGTTTCCTCTAGAAATAATTTTGTAAAGAGGATATACCA 180

Db 121 AGACCACACCGGTTTCCTCTAGAAATAATTTTGTAAAGAGGATATACCA 180

121 AGACCAACCGGTTCCCTCTAGAAATAATTTTGTATTAACTTTAAGAGGAGATATACCA 180
181 TGATTGCTGGACCTGAGTGGCTGTAGACGGTCCATCTCTACACACGCCAATTTGGTTG 240
181 TGATTGCTGGACCTGAGTGGCTGTAGACGGTCCATCTCTACACACGCCAATTTGGTTG 240
241 TTAGCGTTGCTGGTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTTAAATTTTGA 300
241 TTAGCGTTGCTGGTACTGTGTAGGGGACGAATCAAGACATTAAGTCTTTAAATTTTGA 300
301 TCGATTTAACTACGACCTGCTCATGGAGGAAGACAGACGAGCAAGGCTTTAAGTCCAA 360
301 TCGATTTAACTACGACCTGCTCATGGAGGAAGACAGACGAGCAAGGCTTTAAGTCCAA 360
361 CAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATAACTTTGAGAAAGCTGACTT 420
361 CAAAACCAATTTGCTACTGATAGTGGCGGATGTCACATAACTTTGAGAAAGCTGACTT 420
421 TAAAGGCTATTCAAGAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTCTTTGAG 480
421 TAAAGGCTATTCAAGAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTCTTTGAG 480
481 TCATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTCTATTG 540
481 TCATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGCGCAAGGCTCTATTG 540
541 ACAAGATGGTTCGGTAACTTTGCGGACCCCAACCTGTCAGAAATTTTGTCTAAGCG 600
541 ACAAGATGGTTCGGTAACTTTGCGGACCCCAACCTGTCAGAAATTTTGTCTAAGCG 600
601 ATGTGCGGTTAGACCATATAGAAACCAATACAAACCAAGCAAGCAATCTGTTGATG 660
601 ATGTGCGGTTAGACCATATAGAAACCAATACAAACCAAGCAAGCAATCTGTTGATG 660
661 TGGATATATCTGTAAGTTTACTCCCTTAAACCTGATGACCAATTTGACACCAAGGCT 720
661 TGGATATATCTGTAAGTTTACTCCCTTAAACCTGATGACCAATTTGACACCAAGGCT 720
721 AAGATATCTGTAAGTTTACTGATGATCGGTGACACCAATTTGACACCAAGGCT 780
721 AAGATATCTGTAAGTTTACTGATGATCGGTGACACCAATTTGACACCAAGGCT 780
781 TAGCTCAGCAACCAAGTTTAAACCAACCAACCAAGGCTATAGGATTTAAGACG 840
781 TAGCTCAGCAACCAAGTTTAAACCAACCAACCAAGGCTATAGGATTTAAGACG 840
841 ACTCTCAATCTGCTACTGATGATGATTTTCCGTACGATTTTACCAATGGATCAAG 900
841 ACTCTCAATCTGCTACTGATGATGATTTTCCGTACGATTTTACCAATGGATCAAG 900
901 AGTTTACTTACCGTGTAAATTCGGGAAACAGCTTTATAGGATCAATAAAAAATCTG 960
901 AGTTTACTTACCGTGTAAATTCGGGAAACAGCTTTATAGGATCAATAAAAAATCTG 960
961 TGAATGAAGAAATAAACCACTGATGATCTGAGAAATTTAAGTCTTTAAAAAG 1020
961 TGAATGAAGAAATAAACCACTGATGATCTGAGAAATTTAAGTCTTTAAAAAG 1020
1021 GGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAATCTTTTACCATCAAT 1080
1021 GGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAATCTTTTACCATCAAT 1080
1081 TTGATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTTTAAAGCTAGCGAG 1140
1081 TTGATGTCGATACCAACGAATTTGCTAAAAGTGAGCAGCTTTAAAGCTAGCGAG 1140
1141 ACTTAGATCTCAGAGATTTATAGATCTTCTGAGTAAGGCTAAATCTTCTCAACAAT 1200
1141 ACTTAGATCTCAGAGATTTATAGATCTTCTGAGTAAGGCTAAATCTTCTCAACAAT 1200
1201 TCGATGCTTTTGGTATTGACTATACCTTAACCTGGAAGCTAGAGGATATACGATG 1260

1201 TCGATGCTTTTGGTATTATGAGCTATATACCTTAACTGGAAGTAGAGATAATCACGATG 1260
1261 ACACCAACCGTATCATAAACCGTTTATATGCGCAAGCCGAGAGGAGAGAACTCTAGCT 1320
1261 ACACCAACCGTATCATAAACCGTTTATATGCGCAAGCCGAGAGGAGAGAACTCTAGCT 1320
1321 ATCATTTAGCCGCTGGTGGTTCAGGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTG 1380
1321 ATCATTTAGCCGCTGGTGGTTCAGGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTG 1380
1381 TCAAGTCAAGCAAGCCCGGTTGTTATGACAAATGGTTCAGCCCCAGTCCCGGTG 1440
1381 TCAAGTCAAGCAAGCCCGGTTGTTATGACAAATGGTTCAGCCCCAGTCCCGGTG 1440
1441 GCGAGCGGCTACCTAGGTAAATGGTTCAGGCGCAGCAAAATGGTTTCAGCCCCAGTCC 1500
1441 GCGAGCGGCTACCTAGGTAAATGGTTCAGGCGCAGCAAAATGGTTTCAGCCCCAGTCC 1500
1501 TTAACCTGCGAAAGCTAAACCTGGAAGCTGAAGAGCTTCTTTTGAACAAGTACACTGG 1560
1501 TTAACCTGCGAAAGCTAAACCTGGAAGCTGAAGAGCTTCTTTTGAACAAGTACACTGG 1560
1561 CTTCACGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
1561 CTTCACGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
1621 GCATCGGGCTGGGCGAGGAGGAATAAGCTGTACCATCTAA 1661
1621 GCATCGGGCTGGGCGAGGAGGAATAAGCTGTACCATCTAA 1661

RESULT 2

US-09-940-235-5
; Sequence 5, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5

Query Match 77.2%; Score 1282.2; DB 3; Length 1377;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1284; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 51 GCACCCGTTGGCGAGGACCAACCGCTGCCGAGATCTCGATCCCGGAAATTAATACCACT 110
DB 1 GCACCCGTTGGCGAGGACCAACCGCTGCCGAGATCTCGATCCCGGAAATTAATACCACT 60
QY 111 CACTATAGGAGAGACCAACCGGTTTCCCTCTAGAAATAATTTTGTCTTAACTTTAAGAGG 170
DB 61 CACTATAGGAGAGACCAACCGGTTTCCCTCTAGAAATAATTTTGTCTTAACTTTAAGAGG 120

Db	781	AAGCTTATAGGATCAATAAAAAATCTGGTCTGAAATGAGAAATAAACACACTGACCTGA	840
Qy	991	TCTCTGAGAAATATTACGTCCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA	1050
Db	841	TCTCTGAGAAATATTACGTCCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA	900
Qy	1051	GTCACTTGAACCTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTAAAAA	1110
Db	901	GTCACTTGAACCTGTTCCACCATCAAAATACGTTGATGTCGATACCAACGAATTCGTAAAAA	960
Qy	1111	GTGAGCAGCTCTTAAACAGCTAGCGAAACGTAACCTTAGACTTCAGAGATTTATACGATCCTC	1170
Db	961	GTGAGCAGCTCTTAAACAGCTAGCGAAACGTAACCTTAGACTTCAGAGATTTATACGATCCTC	1020
Qy	1171	GTGATTAAGGCTAAACTACTCTACACAACTCTCGATGCTTTGGTATTATGGACTATACCT	1230
Db	1021	GTGATTAAGGCTAAACTACTCTACACAACTCTCGATGCTTTGGTATTATGGACTATACCT	1080
Qy	1231	TAACTGGAAAAAGTAGAGGATTAATCACGATGACACCAACCGTATCATAAACCGTTTATATGG	1290
Db	1081	TAACTGGAAAAAGTAGAGGATTAATCACGATGACACCAACCGTATCATAAACCGTTTATATGG	1140
Qy	1291	GCAAGCGACCCGAAGGAGAGAAATGCTAGCTATCATTTTAGCCGGTGGTGGTCAGGGCGACG	1350
Db	1141	GCAAGCGACCCGAAGGAGAGAAATGCTAGCTACCATTTAGCTGGTGGTGGCCAGGCGCAAC	1200
Qy	1351	AAATGGTTC	1359
Db	1201	AGATTGTAC	1209

RESULT 5

```

US-09-940-235-12
; Sequence 12, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chaic
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Subdaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-12

```

	Query Match	70.2%	Score 1165.8	DB 3	Length 2096
	Best Local Similarity	99.4%	Prod. NO. 0		
	Matches 1170	Conservative	0	Mismatches 7	Indels 0
					Gaps 0
Qy	183	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATTGGTTGTT	242		
Db	588	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATTGGTTGTT	647		

Qy	243	AGCGTTGCTGGTA	CTGTTTGAAGGGAGCGAATCAAGACATTAAGTCTTTAAATTTTTTTGAAATC	307
Db	648	AGCGTTGCTGGTACT	TGTTTGAGGGAGCGAATCAAGACATTAAGTCTTTAAATTTTTTTGAAATC	707
Qy	303	GATCTAAACATACGACCTGCT	CATGAGGGAAGACAGAGCAAGGCTTAAGTCCAAATCA	362
Db	708	GATCTAAACATACGACCTGCT	CATGAGGGAAGACAGAGCAAGGCTTAAGTCCAAATCA	767
Qy	363	AAACCATTTTGCTACT	GTGATAGTGGGCGCATGTCCACATAAACTTTGAGAAAAGCTGACTTACTA	422
Db	768	AAACCATTTTGCTACT	GTGATAGTGGGCGCATGTCCACATAAACTTTGAGAAAAGCTGACTTACTA	827
Qy	423	AAGGCTATTCAAGAAACAAT	TGATTCGTAAAGCTCCACAGTAAACGACGACTACTTTGAGGTC	482
Db	828	AAGGCTATTCAAGAAACAAT	TGATTCGTAAAGCTCCACAGTAAACGACGACTACTTTGAGGTC	887
Qy	483	ATTGATTTTGCAGGATGCA	ACCAATTAACCTGATCGAAACGCAAGGCTACTTTGCTGAC	542
Db	888	ATTGATTTTGCAGGATGCA	ACCAATTAACCTGATCGAAACGCAAGGCTACTTTGCTGAC	947
Qy	543	AAAGATGGTTCCGTAAACCT	TGCCCGACCCAACTCTGTCCAAAGAAATTTTGTCTAAGCGGACAT	602
Db	948	AAAGATGGTTCCGTAAACCT	TGCCCGACCCAACTCTGTCCAAAGAAATTTTGTCTAAGCGGACAT	1007
Qy	603	GTGGCGGTTAGACCATATA	TAAGAAAGAAACCAATAACAAACGAGGAAATCTGTTGATGTG	662
Db	1008	GTGGCGGTTAGACCATATA	TAAGAAAGAAACCAATAACAAACGAGGAAATCTGTTGATGTG	1067
Qy	663	GAATATACGTACAGTTTACT	CCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCTCAA	722
Db	1068	GAATATACGTACAGTTTACT	CCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCTCAA	1127
Qy	723	GATCTAAGCTATTGAAACCA	CTAGCTATCGGTGACACCATCACTCTCAAGAAATTACTA	782
Db	1128	GATCTAAGCTATTGAAACCA	CTAGCTATCGGTGACACCATCACTCTCAAGAAATTACTA	1187
Qy	783	GCTCAAGCAAAAGCATTTT	TAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC	842
Db	1188	GCTCAAGCAAAAGCATTTT	TAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC	1247
Qy	843	TCCTCAATCGTCACATGACA	ATGACATTTTCGTCAGGATTTTACCAATGGATCAAGAG	902
Db	1248	TCCTCAATCGTCACATGACA	ATGACATTTTCGTCAGGATTTTACCAATGGATCAAGAG	1307
Qy	903	TTTACTTACCGTGTTAAAA	ATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
Db	1308	TTTACTTACCGTGTTAAAA	ATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	1367
Qy	963	AATGAAGAAATAAACAAC	ACTGACCTGTCTCGAGAAATATTACGTCCTTTAAAAAAGGG	1022
Db	1368	AATGAAGAAATAAACAAC	ACTGACCTGTCTCGAGAAATATTACGTCCTTTAAAAAAGGG	1427
Qy	1023	GAAGAAGCGGTATGATCC	CTTTTGATCGGAGTCACATTGAAAATGTTTCAAGATCAAAATAGCTT	1082
Db	1428	GAAGAAGCGGTATGATCC	CTTTTGATCGGAGTCACATTGAAAATGTTTCAAGATCAAAATAGCTT	1487
Qy	1083	GATGTGATACCAACGAAT	TGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC	1142
Db	1488	GATGTGATACCAACGAAT	TGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGCGAACGTAAC	1547
Qy	1143	TTAGACTTTCAGAGATTT	ATACGATCTCTCGTGAATAGGCTAAACTACTCTACAACAATCTC	1202
Db	1548	TTAGACTTTCAGAGATTT	ATACGATCTCTCGTGAATAGGCTAAACTACTCTACAACAATCTC	1607
Qy	1203	GATGCTTTTGGTATTTAT	GGACTATACCTTAACTGGAAGAGTGAAGGATAATACGATGAC	1262
Db	1608	GATGCTTTTGGTATTTAT	GGACTATACCTTAACTGGAAGAGTGAAGGATAATACGATGAC	1667
Qy	1263	ACCAACCGTATCATAAAC	CGTTTATATCGGCAAGCGACCGAAGGAGAGATGCTAGCTAT	1322
Db	1668	ACCAACCGTATCATAAAC	CGTTTATATCGGCAAGCGACCGAAGGAGAGATGCTAGCTAT	1727
Qy	1323	CATTTAGCCGCTGGTGT	TCAGGCGCAGCAATGGTTC	1359


```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1242)
US-09-940-235-1

Query Match          69.2%; Score 1148.6; DB 3; Length 1245;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGGTCCATCTGTCAACAACAGCCAAATGGTTGTT 242
Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGGTCCATCTGTCAACAACAGCCAAATGGTTGTT 60

Qy 243 AGCGTTGCTGGTACTCTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 61 AGCGTTGCTGGTACTCTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGGAAATC 120

Qy 303 GATCTAAACATCAGCACTGCTCATGAGGAGAAACAGAGCAAGCTTTAAGTCCAAATCA 362
Db 121 GATCTAAACATCAGCACTGCTCATGAGGAGAAACAGAGCAAGCTTTAAGTCCAAATCA 180

Qy 363 AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAACTTCGAGAAAGCTGACTACTA 422
Db 181 AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAACTTCGAGAAAGCTGACTACTA 240

Qy 423 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 482
Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 300

Qy 483 ATTGATTTTGAAGCGATGGAACCATTAATGATCGGAAACGGGAAGGTCTACTTTGCTGAC 542
Db 301 ATTGATTTTGAAGCGATGGAACCATTAATGATCGGAAACGGGAAGGTCTACTTTGCTGAC 360

Qy 543 AAAGATGGTTCGGTAACTTTGCGGACCCCACTGTCACAGAAATTTTGTCTAAGCGGCAT 602
Db 361 AAAGATGGTTCGGTAACTTTGCGGACCCCACTGTCACAGAAATTTTGTCTAAGCGGCAT 420

Qy 603 GTGCGGTTAGACCATATATAAGAAAAACCAATAACCAAGCGAAATCTGTGTGATGTG 662
Db 421 GTGCGGTTAGACCATATATAAGAAAAACCAATAACCAAGCGAAATCTGTGTGATGTG 480

Qy 663 GAATATAGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAAGCAGGTCTCAA 722
Db 481 GAATATAGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAAGCAGGTCTCAA 540

Qy 723 GATCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
Db 541 GATCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 600

Qy 783 GCTCAAGCACAAGCAATTTAAACAAAAACCCACCGGCTATACGATTTTATGAACGTGAC 842
Db 601 GCTCAAGCACAAGCAATTTAAACAAAAACCCACCGGCTATACGATTTTATGAACGTGAC 660

Qy 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720

Qy 903 TTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCTG 962
Db 721 TTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCTG 780

Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATACCTCTTAAAAAGG 1022
Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATACCTCTTAAAAAGG 840

Qy 1023 GAAAAGCCGATATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCAATCAATACGTT 1082
Db 841 GAAAAGCCGATATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCAATCAATACGTT 900
```

```

Qy 1083 GATCTGATACCAACGAATTTGCTAAAAAGTGAAGAGTCTTAAACAGCTAGCGAACTAAC 1142
Db 901 GATCTGATACCAACGAATTTGCTAAAAAGTGAAGAGTCTTAAACAGCTAGCGAACTAAC 960

Qy 1143 TTAGACTTCAGAGATTTTATATAGATCCCTGCTGATAGGCTAAACTCTCTACAACTTC 1202
Db 961 TTAGACTTCAGAGATTTTATATAGATCCCTGCTGATAGGCTAAACTCTCTACAACTTC 1020

Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATTAATCAGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATTAATCAGATGAC 1080

Qy 1263 ACCAACCGTATCATAAACGGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 1081 ACCAACCGTATCATAAACGGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1140

Qy 1323 CATTTAGCCCGGTGGT 1337
Db 1141 CATTTAGCCCTATGAT 1155

RESULT 8
US-10-474-792-657
; Sequence 657, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 657
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-10-474-792-657

Query Match          59.9%; Score 995; DB 8; Length 1323;
Best Local Similarity 91.3%; Pred. No. 3.6e-274;
Matches 1055; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGGTCCATCTGTCAACAACAGCCAAATGGTTGTT 242
Db 79 ATTGCTGGGTATGGATGGCTACCGACCGTCCACCTATCAATAACGCCAGTTAGTTGTT 138

Qy 243 AGCGTTGCTGGTACTCTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 139 AGTATGCGCGGTATCGTTGAAAGGTACCGATAAAAAAGTTTTTATAAAATTTTTTGGAAATC 198

Qy 303 GATCTAATCATCAGACCTGCTCATGAGGAGAAAGCAGAGCAAGGCTTTAAGTCCAAATCA 362
Db 199 GATCTAATCATCAGACCTGCTCATGAGGAGAAAGCAGAGCAAGGCTTTAAGTCCAAATCA 258

Qy 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAAGAGCTGACTTACTA 422
Db 259 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAAGAGCTGACTTACTA 318

Qy 423 AAGGCTATTCAAGAAACAAATTTGATCGCTTAACGCTCCACAGTAACCGACGACTACTTTGAGGTC 482
Db 319 AAGGCTATTCAAGAAACAAATTTGATCGCTTAACGCTCCACAGTAACCGACGACTACTTTGAGGTC 378

Qy 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
Db 379 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 438

Qy 543 AAAGATGGTTCGGTAACTTTGCGGACCCCACTGTCACAGAAATTTTGTCTAAGCGGCAT 602
Db 379 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 438
```


Db 439 AAGATGGTTCGGTAACTCTGGCGACCCCACTCTGCCAAGAAATTTTGTGTTAAAGGACAT 498
Qy 603 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 499 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 558
Qy 663 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 722
Db 559 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 618
Qy 723 GATACATAGCTATTGAACACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTAATA 782
Db 619 GATACATAGCTATTGAACACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTAATA 678
Qy 783 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGCCAGGCTATACGATTTTATGATGACGTGAC 842
Db 679 GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGCCAGGCTATACGATTTTATGATGACGTGAC 738
Qy 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 739 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 798
Qy 903 TTTACTTACCGTTTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAATCTGGTCTG 962
Db 799 TTTACTTACCATGTCAAAAATCGGGAACAAAGCTTATGAGATCAATCTTAAAAACAGGTATT 858
Qy 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 859 AAGAAAAAACAAGCAACTGATCTGAGTCTCTGAGAAATATTACGTCCTTAAAAAGGG 918
Qy 1023 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACCATCAATACGTT 1082
Db 919 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACCATCAATACGTT 978
Qy 1083 GATGTCGATACCAACGAATGTGTAAGTAGACAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 979 GATGTCGATACCAACGAATGTGTAAGTAGACAGCTCTTAAACAGCTAGCGAACGTAAC 1038
Qy 1143 TTAGACTTCAGAGATTATACGATCTCTGATAGAGCTTAAACTACTCTACACATCTC 1202
Db 1039 TTAGACTTCAGAGATTATACGATCTCTGATAGAGCTTAAACTACTCTACACATCTC 1098
Qy 1203 GATGCTTTGTTATGATGACATATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1262
Db 1099 GATGCTTTGTTATGATGACATATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1158
Qy 1263 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1322
Db 1159 AATAATCGTGTGTTACAGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1218
Qy 1323 CATTTAGCGGTGGT 1337
Db 1219 CATTTAGCTTATGAT 1233

RESULT 9
US-10-236-392-1

; Sequence 1, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladamir

Query Match 19.1%; Score 316.8; DB 7; Length 6988;
Best Local Similarity 99.4%; Pred. No. 4e-79; Mismatches 0; Indels 0; Gaps 0;
Matches 318; Conservative 0

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGTGTGTCAGTCAAAAGCAAGCCG 1398
Db 12 GGCAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGTGTGTCAGTCAAAAGCAAGCCG 71

Qy 1399 GTTGTATGCAATGGAAACACTATCAGATAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGCAATGGAAACACTATCAGATAATCAACAGTGGAGCGGACCTACCTAG 131

Qy 1459 GTAAATGTTGGTTTGTACTTGTATCGAGAACCCAGGTTTAACTGCCAAAAGTAAAC 1518
Db 132 GTAAATGTTGGTTTGTACTTGTATCGAGAACCCAGGTTTAACTGCCAAAAGTAAAC 191

Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAAACACTTACCGAGTGGGTGACA 251

; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRoche, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Maiyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard A
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-442A
; CURRENT APPLICATION NUMBER: US/10/236,392
; CURRENT FILING DATE: 2002-09-06
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US60/390,155
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US09/635,949
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US60/318,765
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/357,303
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US60/367,753
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US60/369,479
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US09/659,634
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US60/318,120
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US60/318,130
; PRIOR FILING DATE: 2001-09-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 1
; LENGTH: 6988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26)..(6986)
; US-10-236-392-1

QY 1579 CTTATGAGGCTCTAAGACTCCATGATCTGGACGTGTACCTGCGATCGGGCTGGGGCGAG 1638
Db 252 CTTATGAGGCTCTAAGACTCCATGATCTGGACGTGTACCTGCGATCGGGCTGGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 312 GGAGATAAGCTGTACCATC 331

RESULT 10

US-10-831-704-38
; Sequence 38, Application US/10831704
; Publication No. US20050100931A1
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND
; TITLE OF INVENTION: ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/10/831,704
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US/10/155,653
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US/09/220,132
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-831-704-38

Query Match 19.1%; Score 316.8; DB 9; Length 7679;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1339 GTCAGGCCAGCAAAATGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAATGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 71
QY 1399 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 131
QY 1459 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1518
Db 132 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 191

QY 1519 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 251
QY 1579 CTTATGAGCGTCTAAGACTCCATGATCTGGGACTGTGACATCGGGCTGGGGCGAG 1638
Db 252 CTTATGAGCGTCTAAGACTCCATGATCTGGGACTGTGACATCGGGCTGGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 312 GGAGATAAGCTGTACCATC 331

RESULT 11

US-09-964-824A-574
; Sequence 574, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrihan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 574
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-574

Query Match 19.1%; Score 316.8; DB 3; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1339 GTCAGGCCAGCAAAATGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAAAATGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCG 71

QY 1399 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 131
QY 1459 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1518
Db 132 GTTGTATGACATCGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 191
QY 1519 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 251

QY 1579 CTTATGAGCGTCTAAGACTCCATGATCTGGGACTGTGACATCGGGCTGGGGCGAG 1638
Db 252 CTTATGAGCGTCTAAGACTCCATGATCTGGGACTGTGACATCGGGCTGGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 312 GGAGATAAGCTGTACCATC 331

RESULT 12

US-10-171-311-63
; Sequence 63, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-63

Query Match          19.1%; Score 316.8; DB 5; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCACTGTCAGTCCCGGTGCTGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAAAATGGTTTCAGCCCACTGTCAGTCCCGGTGCTGTCAAGCAAGCCCG 71

Qy 1399 GTTGTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 131

Qy 1459 GTAATGTGTTGTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 1518
Db 132 GTAATGTGTTGTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 191

Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAAACACTTACCGAGTGGGTGACA 251

Qy 1579 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 1638
Db 252 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 311

Qy 1639 GGAGATAAAGCTGTACCATC 1658
Db 312 GGAGATAAAGCTGTACCATC 331

RESULT 14
US-10-374-979-75
; Sequence 75, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-75

Query Match          19.1%; Score 316.8; DB 6; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCACTGTCAGTCCCGGTGCTGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAAAATGGTTTCAGCCCACTGTCAGTCCCGGTGCTGTCAAGCAAGCCCG 71

Qy 1399 GTTGTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 131

Qy 1459 GTAATGTGTTGTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 1518
Db 132 GTAATGTGTTGTTGTACTTGTATGGAGGAAGCCGAGGTTTAACTCGGAAAGTAAAC 191

Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAAACACTTACCGAGTGGGTGACA 1578
Db 192 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAAACACTTACCGAGTGGGTGACA 251

Qy 1579 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 1638
Db 252 CTTATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 311

Qy 1639 GGAGATAAAGCTGTACCATC 1658
Db 312 GGAGATAAAGCTGTACCATC 331

RESULT 13
US-10-236-031B-69
; Sequence 69, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-031B-69

Query Match          19.1%; Score 316.8; DB 6; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCACTGTCAGTCCCGGTGCTGTCAAGCAAGCCCG 1398
Db 12 GGCAGGCTCAGCAAAATGGTTTCAGCCCACTGTCAGTCCCGGTGCTGTCAAGCAAGCCCG 71

Qy 1399 GTTGTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 72 GTTGTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 131
```

RESULT 15
US-10-182-936A-75
; Sequence 75, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-75

Query Match 19.1%; Score 316.8; DB 7; Length 7680;
Best Local Similarity 99.4%; Pred. No. 4.2e-79;
Matches 318; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1339 GTGAGCGGAGCAAAATGGTTTCAGCCCGAGTCCCGGGTGTTCAGTCAAAAGCAAGCCCG 1398
Db |||||||
12 GGCAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGGTGTTCAGTCAAAAGCAAGCCCG 71
QY 1399 GTTGTATGACATGGAAACACTATCAGATAATCAACAGTGGAGGGGACCTACCTAG 1458
Db |||||||
72 GTTGTATGACATGGAAACACTATCAGATAATCAACAGTGGAGGGGACCTACCTAG 131
QY 1459 GTAATGTGTGGTTGTCTTGTATGGAGAGCCGAGGTTTTTAACCTGCGAAAGTAAAC 1518
Db |||||||
132 GTAATGTGTGGTTGTCTTGTATGGAGAGCCGAGGTTTTTAACCTGCGAAAGTAAAC 191
QY 1519 CTGAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578
Db |||||||
192 CTGAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 251
QY 1579 CTTATGAGGCTCTAAAGACTCCATGATCTGGGACTGTACTGTGATCGGGGCTGGGCGAG 1638
Db |||||||
252 CTTATGAGGCTCTAAAGACTCCATGATCTGGGACTGTACTGTGATCGGGGCTGGGCGAG 311
QY 1639 GGAGATAAGCTGTACCATC 1658
Db |||||||
312 GGAGATAAGCTGTACCATC 331

Search completed: February 1, 2006, 14:21:19
Job time : 1373.44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 18:17:59 ; Search time 307.097 Seconds
(without alignments)
4492.841 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaaccgccagcctagcc.....gaataagctgaccatctaa 1661

Scoring table: IDENTITY NUC
Gap 10.0 , Gapext 1.0

Searched: 6059551 seqs, 415333918 residues

Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	100.0	1661	7	US-10-631-558-10
2	1282.2	77.2	1377	7	US-10-631-558-5
3	1219.4	73.4	1327	7	US-10-631-558-6
4	1185	71.3	1541	7	US-10-631-558-9
5	1165.8	70.2	2096	7	US-10-631-558-12
6	1150.2	69.2	1782	7	US-10-631-558-11
7	1148.6	69.2	1245	7	US-10-631-558-1
8	316.4	19.0	777	7	US-10-631-558-3
9	312	18.8	8232	7	US-10-821-234-693
10	311.6	18.8	2443	7	US-10-995-561-114
11	311.6	18.8	2488	7	US-10-995-561-105
12	311.6	18.8	6510	7	US-10-995-561-112
13	311.6	18.8	7823	7	US-10-995-561-117
14	311.6	18.8	7848	7	US-10-995-561-111
15	311.6	18.8	7935	7	US-10-995-561-113
16	311.6	18.8	7959	7	US-10-995-561-108
17	311.6	18.8	8013	7	US-10-995-561-101
18	311.6	18.8	8155	7	US-10-995-561-116
19	311.6	18.8	8226	7	US-10-995-561-107
20	311.6	18.8	8278	7	US-10-995-561-106
21	311.6	18.8	8332	7	US-10-995-561-110
22	311.6	18.8	8371	7	US-10-995-561-109

23	255	15.4	8404	8	US-11-136-527-2446	Sequence 2446, Ap
24	135.8	8.2	87672	7	US-10-995-561-13237	Sequence 13237, A
25	133.6	8.0	201	7	US-10-995-561-4446	Sequence 4446, Ap
26	133.6	8.0	201	7	US-10-995-561-4450	Sequence 4450, Ap
27	133.6	8.0	201	7	US-10-995-561-4468	Sequence 4468, Ap
28	133.6	8.0	201	7	US-10-995-561-4486	Sequence 4486, Ap
29	133.6	8.0	201	7	US-10-995-561-4503	Sequence 4503, Ap
30	133.6	8.0	201	7	US-10-995-561-4522	Sequence 4522, Ap
31	133.6	8.0	201	7	US-10-995-561-4541	Sequence 4541, Ap
32	133.6	8.0	201	7	US-10-995-561-4560	Sequence 4560, Ap
33	133.6	8.0	201	7	US-10-995-561-4575	Sequence 4575, Ap
34	133.6	8.0	201	7	US-10-995-561-4592	Sequence 4592, Ap
35	133.6	8.0	201	7	US-10-995-561-4596	Sequence 4596, Ap
36	133.6	8.0	201	7	US-10-995-561-4621	Sequence 4621, Ap
37	133.6	8.0	201	7	US-10-995-561-4640	Sequence 4640, Ap
38	133.6	8.0	201	7	US-10-995-561-27119	Sequence 27119, A
39	99	6.0	4974	8	US-11-094-586-17	Sequence 17, Appl
40	99	6.0	4974	8	US-11-076-733-86	Sequence 86, Appl
41	99	6.0	8966	8	US-11-076-733-87	Sequence 87, Appl
42	98	5.9	227	8	US-11-137-395-9	Sequence 9, Appl
43	98	5.9	5303	8	US-11-149-403-34	Sequence 34, Appl
44	98	5.9	7085	8	US-11-149-403-33	Sequence 33, Appl
45	98	5.9	8031	8	US-11-149-403-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-10-631-558-10
; Sequence 10, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631.558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940.235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-10

Query Match 100.0%; Score 1661; DB 7; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAACCCCGCAGCCTAGCCGGTCTTCAACGACGAGGACGATCATGCGACCCGTTGG 60
Db 1 GCAACCCCGCAGCCTAGCCGGTCTTCAACGACGAGGACGATCATGCGACCCGTTGG 60
Qy 61 CCAGAGCCCAACCGTCCCGAGATCGATCCCGGAATTAAATACGACTCACTATAGGG 120

Db 61 CCAGGACCCAAACGCTGCCGAGATCTCGATCCCGCGAATTAATACGACTCACTATAGGG 120
 Qy 121 AGACACCAACGCTTCCCTCTAGAAATAATTTTGTGTTTAACTTTAAAGAGGAGATATACCA 180
 Db 121 AGACCAACGCTTCCCTCTAGAAATAATTTTGTGTTTAACTTTAAAGAGGAGATATACCA 180
 Qy 181 TGAATGCTGAGCTGAGTGGCTGCTAGAGCGTCCATCTGTCAACACAGCCAAATTTGGTTG 240
 Db 181 TGAATGCTGAGCTGAGTGGCTGCTAGAGCGTCCATCTGTCAACACAGCCAAATTTGGTTG 240
 Qy 241 TTAGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db 241 TTAGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Qy 301 TCGATCTACATCACGACCTGCTCATGGAGGAAGACAGAGCGCTTAAGTCCCAAAAT 360
 Db 301 TCGATCTACATCACGACCTGCTCATGGAGGAAGACAGAGCGCTTAAGTCCCAAAAT 360
 Qy 361 CAAAACCAATTTGCTACTGATAGTGGCGCATGTGCATATAACTTTGAGAAAGCTGACTTAC 420
 Db 361 CAAAACCAATTTGCTACTGATAGTGGCGCATGTGCATATAACTTTGAGAAAGCTGACTTAC 420
 Qy 421 TAAAGCTATTCAAGAACCAATTTGATCGCTAAACCTCAAGTAAACGACGACTCTTTGAGG 480
 Db 421 TAAAGCTATTCAAGAACCAATTTGATCGCTAAACCTCAAGTAAACGACGACTCTTTGAGG 480
 Qy 481 TCAATGATTTTCAAGCGATGCAACCAATTTACTGATCGAAGCGGCAAGGCTTACTTTGCTG 540
 Db 481 TCAATGATTTTCAAGCGATGCAACCAATTTACTGATCGAAGCGGCAAGGCTTACTTTGCTG 540
 Qy 541 ACAAAGATGTTTGGGTAACTTGGCGACCAACCTGTCCAAAGAAATTTTCTTAAGCGGAC 600
 Db 541 ACAAAGATGTTTGGGTAACTTGGCGACCAACCTGTCCAAAGAAATTTTCTTAAGCGGAC 600
 Qy 601 ATGTGCGGCTTACCATATAAGAAACCAATACAAACCAAGCGAAATCTTGTGATG 660
 Db 601 ATGTGCGGCTTACCATATAAGAAACCAATACAAACCAAGCGAAATCTTGTGATG 660
 Qy 661 TGGAAATATCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGACGAGTCTCA 720
 Db 661 TGGAAATATCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGACGAGTCTCA 720
 Qy 721 AAGTACTAAGCTATTGAAACCACTAGCTATGCTGTCACACCATCTCAAGAAATAC 780
 Db 721 AAGTACTAAGCTATTGAAACCACTAGCTATGCTGTCACACCATCTCAAGAAATAC 780
 Qy 781 TAGCTCAAGCACAAGCAATTTTAAACCAAAACCAAGCTATAGGATTTATGAACGTG 840
 Db 781 TAGCTCAAGCACAAGCAATTTTAAACCAAAACCAAGCTATAGGATTTATGAACGTG 840
 Qy 841 ACTCCTCAATCTGCTATGACAAATGACATTTTCCGTAGGATTTTACCAATGGATCAAG 900
 Db 841 ACTCCTCAATCTGCTATGACAAATGACATTTTCCGTAGGATTTTACCAATGGATCAAG 900
 Qy 901 AGTTTACTTACCGTGTAAATATCGGAAACAGCTTATAGGATCAATTAATAATCTGGTC 960
 Db 901 AGTTTACTTACCGTGTAAATATCGGAAACAGCTTATAGGATCAATTAATAATCTGGTC 960
 Qy 961 TGAATGAAGAAATAAACACACTGCTGATCTCTGAGAAATATAGCTCTTAAATAAG 1020
 Db 961 TGAATGAAGAAATAAACACACTGCTGATCTCTGAGAAATATAGCTCTTAAATAAG 1020
 Qy 1021 GGGAAAGCGGTATGATCCCTTGTATCGGAGTCACTTGAACCTGTTTACCAATCAATACG 1080
 Db 1021 GGGAAAGCGGTATGATCCCTTGTATCGGAGTCACTTGAACCTGTTTACCAATCAATACG 1080
 Qy 1081 TTGATGTCGATACCAAGAAATGCTTAAAGTGAAGGAGGAGCTTTAAAGCTAGGAAAG 1140
 Db 1081 TTGATGTCGATACCAAGAAATGCTTAAAGTGAAGGAGGAGCTTTAAAGCTAGGAAAG 1140
 Qy 1141 ACTTAGACTTTCAGAGATTTATAGGATCTCTGATTAAGGCTTAACTTCTACAAATC 1200
 Db 1141 ACTTAGACTTTCAGAGATTTATAGGATCTCTGATTAAGGCTTAACTTCTACAAATC 1200

Qy 1201 TCGATGCTTTTGGTATTATGAGCTATACCTTAACCTGAAAAGTAGAGGATTAATCAGCATG 1260
 Db 1201 TCGATGCTTTTGGTATTATGAGCTATACCTTAACCTGAAAAGTAGAGGATTAATCAGCATG 1260
 Qy 1261 ACACCAACCGGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCT 1320
 Db 1261 ACACCAACCGGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGATGCTAGCT 1320
 Qy 1321 ATCATTTAGCCGCTGGTGGTCAGCGCAGCAAAATGGTTTCCGCCAGTCCCGGTGGCTG 1380
 Db 1321 ATCATTTAGCCGCTGGTGGTCAGCGCAGCAAAATGGTTTCCGCCAGTCCCGGTGGCTG 1380
 Qy 1381 TCAGTCAAAAGCAAGCCCGGTTTGTATGACAAATGAAAAACACTATCAGATAAATCAACAGT 1440
 Db 1381 TCAGTCAAAAGCAAGCCCGGTTTGTATGACAAATGAAAAACACTATCAGATAAATCAACAGT 1440
 Qy 1441 GGGAGCGGACCTACTAGGTAAATGTGTGGTTTGTACTTTGTATGGAGGAAGCCGAGGTT 1500
 Db 1441 GGGAGCGGACCTACTAGGTAAATGTGTGGTTTGTACTTTGTATGGAGGAAGCCGAGGTT 1500
 Qy 1501 TTAACCTCGAAAAGTAAACCTTGAAGCTGAAGAGCTTGGTTTGAACAATCACTGGGAACA 1560
 Db 1501 TTAACCTCGAAAAGTAAACCTTGAAGCTGAAGAGCTTGGTTTGAACAATCACTGGGAACA 1560
 Qy 1561 CTTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
 Db 1561 CTTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCT 1620
 Qy 1621 GCATCGGGGCTGGCGAGGAGGAATAAGCTGTACCATCTAA 1661
 Db 1621 GCATCGGGGCTGGCGAGGAGGAATAAGCTGTACCATCTAA 1661

RESULT 2

US-10-631-558-5
 ; Sequence 5, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 1377
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-631-558-5

Query Match 77.2%; Score 1282.2; DB 7; Length 1377;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1284; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 51 GCACCGTGGCCAGGACCCACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACT 110

Db 1 GCACCCGTGCCAGGACCCCAACGCTGCCGAGATCTCGATCCCGGAAATTAATACGACT 60
 QY 111 CACTATAGGAGAGACCAACAGGTTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGG 170
 Db 61 CACTATAGGAGAGACCAACAGGTTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGG 120
 QY 171 AGATATACCATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGC 230
 Db 121 AGATATACCATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGC 180
 QY 231 CAATTGGTCTGTAGCGTTGCTGTGTAAGTGGTGGAGGAGCAATCAAGACATTAGTCTTAAA 290
 Db 181 CAATTGGTCTGTAGCGTTGCTGTGTAAGTGGTGGAGGAGCAATCAAGACATTAGTCTTAAA 240
 QY 291 TTTTGTGAATCGATCTTAACATACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTA 350
 Db 241 TTTTGTGAATCGATCTTAACATACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTA 300
 QY 351 AGTCCAAATCAAAACCAATTTGCTACTGTAGTGGCGGATGTGCATATAAATTGAGAAA 410
 Db 301 AGTCCAAATCAAAACCAATTTGCTACTGTAGTGGCGGATGTGCATATAAATTGAGAAA 360
 QY 411 GCTGACTTACTAAGGCTATTCAAGACCAATTTGATCGTAACTGCTCAAGTAAACGAC 470
 Db 361 GCTGACTTACTAAGGCTATTCAAGACCAATTTGATCGTAACTGCTCAAGTAAACGAC 420
 QY 471 TACTTTGAGTCAATGATTTTGAAGCGATGCAACCAATTAAGTAAAGGCAAGGCTC 530
 Db 421 TACTTTGAGTCAATGATTTTGAAGCGATGCAACCAATTAAGTAAAGGCAAGGCTC 480
 QY 531 TACTTTGCTGACAAAGATGTTTCGTTAACTTGCAGCAACCTGCTCAAGAAATTTTG 590
 Db 481 TACTTTGCTGACAAAGATGTTTCGTTAACTTGCAGCAACCTGCTCAAGAAATTTTG 540
 QY 591 CTAAGCGGACATGTCGCGTTAGACCATATTAAGAAACCAATACAAAACCAAGGAAA 650
 Db 541 CTAAGCGGACATGTCGCGTTAGACCATATTAAGAAACCAATACAAAACCAAGGAAA 600
 QY 651 TCTGTTGATGTGAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGTTCAGA 710
 Db 601 TCTGTTGATGTGAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGTTCAGA 660
 QY 711 CCAGGCTCAAGATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCT 770
 Db 661 CCAGGCTCAAGATATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCT 720
 QY 771 CAAGAAATTAAGTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATT 830
 Db 721 CAAGAAATTAAGTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATT 780
 QY 831 TATGACGTCGATCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCA 890
 Db 781 TATGACGTCGATCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCA 840
 QY 891 ATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAA 950
 Db 841 ATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACCAAGCTTATAGGATCAATAA 900
 QY 951 AAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTC 1010
 Db 901 AAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTC 960
 QY 1011 CTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGAGTCACTTTGAAACTGTTTCAAC 1070
 Db 961 CTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGAGTCACTTTGAAACTGTTTCAAC 1020
 QY 1071 ATCAATAGCTTGATGTCGATACCAAGCAATTTGATTAAGTGGAGGCTCTTAAAGCT 1130
 Db 1021 ATCAATAGCTTGATGTCGATACCAAGCAATTTGATTAAGTGGAGGCTCTTAAAGCT 1080
 QY 1131 AGCGAAGCTAACTTAGACTTACAGAGATTTATAGATCTCTCGTATAGGCTTAAACTACT 1190
 Db 1081 AGCGAAGCTAACTTAGACTTACAGAGATTTATAGATCTCTCGTATAGGCTTAAACTACT 1140

QY 1191 TACAACAATCTCGATCTTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGAT 1250
 Db 1141 TACAACAATCTCGATCTTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGAT 1200
 QY 1251 AATCAGATGACCAACACCGTATCAACCGTTTATATGGCAAGCGACCCGAAGAGAG 1310
 Db 1201 AATCAGATGACCAACACCGTATCAACCGTTTATATGGCAAGCGACCCGAAGAGAG 1260
 QY 1311 AATGCTAGCTATCATTTAGCCGCTGT 1337
 Db 1261 AATGCTAGCTATCATTTAGCCCTATGAT 1287

RESULT 3
 US-10-631-558-6
 ; Sequence 6, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; FILE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ; PRIOR FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1327
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-631-558-6

Query Match 73.4%; Score 1219.4; DB 7; Length 1327;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1226; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 101 TAATACGACTCACTATAGGAGAGACCAACCGTTTCCCTCTAGAAATAATTTGTTTAAAC 160
 Db 1 TAATACGACTCACTATAGGAGAGACCAACCGTTTCCCTCTAGAAATAATTTGTTTAAAC 60
 QY 161 TTTAAGAGGAGATATACCATGATTTCTGGAAGTGGCTGTAGACCGTCAATCTGT 220
 Db 61 TTTAAGAGGAGATATACCATGATTTCTGGAAGTGGCTGTAGACCGTCAATCTGT 120
 QY 221 CAACACAGCCAAATTTGTTAGCTTTAGCTTCTGAGGAGCAATCAAGACAT 280
 Db 121 AAATAACAGCCAAATTTGTTAGCTTTGCTGGTACTGTTGAGGGGAGCAATCAAGACAT 180
 QY 281 TAGTCTTAAATTTTGAATTCGATCAACATCAACGCTCTCATGAGGAGAAAGACAGA 340
 Db 181 TAGTCTTAAATTTTGAATTCGATCAACATCAACGCTCTCATGAGGAGAAAGACAGA 240
 QY 341 GCAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATPAA 400
 Db 241 GCAAGGCTTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCAATPAA 300
 QY 401 ACTTGAGAAAGCTACTTAAAGGCTATTCAAGAACCAATTTGATCGTAACTGCTTAACTCCAC 460


```

Db 301 ACTTGAGAAAGTGACTTAAAGGCTATTCAGAAACAATTGATCGCTAACGTCACAG 360
Qy 461 TAAACAGCACTACTTTGAGGTCAATGATTTTCAAGCGATGCAACCAATTAATGATCGAAA 520
Db 361 TAAACAGCACTACTTTGAGGTCAATGATTTTCAAGCGATGCAACCAATTAATGATCGAAA 420
Qy 521 CGGCAAGGCTACTTTGCTGACAAAGATGGTTCGGTAACTTCGCGACCAACCTGTCCA 580
Db 421 CGGCAAGGCTACTTTGCTGACAAAGATGGTTCGGTAACTTCGCGACCAACCTGTCCA 480
Qy 581 AGAATTTTCTTAAGCGGATGTCGCGTTAGACCAATATAAAGAAACCAATACAAA 640
Db 481 AGAATTTTCTTAAGCGGATGTCGCGTTAGACCAATATAAAGAAACCAATACAAA 540
Qy 641 CCAAGCGAATCTGTTGATGGAATATATCTGACAGTTTACCTCCCTTAAACCTGTATGA 700
Db 541 CCAAGCGAATCTGTTGATGGAATATATCTGACAGTTTACCTCCCTTAAACCTGTATGA 600
Qy 701 CGATTTTCAGACCAAGTCTCAAGATATTAAGCTATTGAAACCACTAGCTATCGGTGACAC 760
Db 601 CGATTTTCAGACCAAGTCTCAAGATATTAAGCTATTGAAACCACTAGCTATCGGTGACAC 660
Qy 761 CATCATCTCAAGAAATTAAGTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGG 820
Db 661 CATCATCTCAAGAAATTAAGTCAAGCAAAAGCATTTTAAACCAAAACCAACCCAGG 720
Qy 821 CTATACGATTTATGACGTCGCTCCTCAATCGTCATGACCAATGACATTTTCCGTAC 880
Db 721 CTATACGATTTATGACGTCGCTCCTCAATCGTCATGACCAATGACATTTTCCGTAC 780
Qy 881 GATTTTACCAATCGGATCAAGAGTTTACTTACCGTGTAAACATCGGGAACCAAGCTATAG 940
Db 781 GATTTTACCAATCGGATCAAGAGTTTACTTACCGTGTAAACATCGGGAACCAAGCTATAG 840
Qy 941 GATCAATAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAA 1000
Db 841 GATCAATAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAA 900
Qy 1001 ATATTACGTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAA 1060
Db 901 ATATTACGTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAA 960
Qy 1061 ACTGTTTACCATCAATACGTTGATGTCGATACCAAGCAATTCCTTAAAGTGCAGAGCT 1120
Db 961 ACTGTTTACCATCAATACGTTGATGTCGATACCAAGCAATTCCTTAAAGTGCAGAGCT 1020
Qy 1121 CTTAAGCAGTAGCGAAGCTTAACTTAGACTTTCAGAGATTTTATACGATCCTCGTGAAGGC 1180
Db 1021 CTTAAGCAGTAGCGAAGCTTAACTTAGACTTTCAGAGATTTTATACGATCCTCGTGAAGGC 1080
Qy 1181 TAACTACTCTACCAATCTCGATGCTTTGATGTTATATGAGTATACCTTAACTGAGAA 1240
Db 1081 TAACTACTCTCTACCAATCTCGATGCTTTGATGTTATATGAGTATACCTTAACTGAGAA 1140
Qy 1241 AGTAGAGGATATACGATGACCAACCGTATCATACCGTTTATATGCGGAGCGACC 1300
Db 1141 AGTAGAGGATATACGATGACCAACCGTATCATACCGTTTATATGCGGAGCGACC 1200
Qy 1301 CGAAGGAGAGATGCTAGCTATCATTTAGCCGGTGGT 1337
Db 1201 CGAAGGAGAGATGCTAGCTATCATTTAGCCGTATGAT 1237

```

RESULT 4

US-10-631-558-9

; Sequence 9, Application US/10631558

; Publication No. US20050260598A1

; GENERAL INFORMATION:

; APPLICANT: Kumar, Rajesh

; APPLICANT: Sahni, Girish

; APPLICANT: Roy, Chait

; APPLICANT: Rajagopal, Kammar

; APPLICANT: Nihalani, Deepak

```

; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
; US-10-631-558-9

```

Query Match 71.3%; Score 1185; DB 7; Length 1541;

Best Local Similarity 98.8%; Pred. No. 0; Mismatches 15; Indels 0; Gaps 0; Matches 1194; Conservative 0;

```

Qy 151 TTTGTTTAACTTTAAGAGGAGATATACCATGATTGCTGGACCTGAGTGGCTGTAGACC 210
Db 1 TTTGTTTAACTTTAAGAGGAGATATACCATGATTGCTGGACCTGAGTGGCTGTAGATC 60
Qy 211 GTCCATCTGTCAACAAGCCAAATGTTGTTAGCGTTGCTGGTACTGTTGAGGAGCGA 270
Db 61 GTCCATCTGTCAACAAGCCAAATGTTGTTAGCGTTGCTGGTACTGTTGAGGAGCGA 120
Qy 271 ATCAAGACATTAAGTCTTAAATTTTGAATCGATCTAAACATCAAGCTGCTCATGAG 330
Db 121 ATCAAGACATTAAGTCTTAAATTTTGAATCGATCTAAACATCAAGCTGCTCATGAG 180
Qy 331 GAAAGACAGAGCAAGGCTTAAAGTCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 390
Db 181 GAAAGACAGAGCAAGGCTTAAAGTCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 240
Qy 391 TGTCACATAAATTTGAGAAAGTGAATTAATAAGGCTATTCAAGAGCTATTCAAGAAACA 450
Db 241 TGTCACATAAATTTGAGAAAGTGAATTAATAAGGCTATTCAAGAGCTATTCAAGAAACA 300
Qy 451 ACCTCCACAGTAAGCAAGCTACTTTGAGGTCATTTGAGGTCATTTGAGGTCATTTGAG 510
Db 301 ACCTCCACAGTAAGCAAGCTACTTTGAGGTCATTTGAGGTCATTTGAGGTCATTTGAG 360
Qy 511 CTGATCGAAACGGCAAGGCTACTTTGCTGCAAAAGATGGTTCGGTAACTTGGCGACCC 570
Db 361 CTGATCGAAACGGCAAGGCTACTTTGCTGCAAAAGATGGTTCGGTAACTTGGCGACCC 420
Qy 571 AACCTGTCCAAAGATTTTGTCTAAGCGGCAATGTCGCGGTTAGACCATATAAGAAAC 630
Db 421 AACCTGTCCAAAGATTTTGTCTAAGCGGCAATGTCGCGGTTAGACCATATAAGAAAC 480
Qy 631 CAATACAAAACCAAGCGAATCTGTTGATGTTGGAATATCTGTTGATGTTGTTTACTCT 690
Db 481 CAATACAAAACCAAGCGAATCTGTTGATGTTGGAATATCTGTTGATGTTTACTCTCT 540
Qy 691 ACCTGATGACGATTTTCAGACCAAGCTCTCAAGATATCTAAAGTATTGAAACACTTAG 750
Db 541 ACCTGATGACGATTTTCAGACCAAGCTCTCAAGATATCTAAAGTATTGAAACACTTAG 600
Qy 751 TCGGTGACCAATCACAATCTCAAGAAATTAAGTCTCAAGCAACAAGCAATTTTAAACAAA 810
Db 601 TCGGTGACCAATCACAATCTCAAGAAATTAAGTCTCAAGCAACAAGCAATTTTAAACAAA 660

```

Qy	811	ACCA	CCAGGCTATACGATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAATGACA	870
Db	661	ACCAC	CCAGGCTATACGATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAATGACA	720
Qy	871	TTTT	CCGTACGATTTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC	930
Db	721	TTTT	CCGTACGATTTTTACCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAAC	780
Qy	931	AAGCTT	ATAGGATCAATAAAAAATCTGGTCTGAAATGAAGAAATAAACAACTGACCTGA	990
Db	781	AAGCTT	ATAGGATCAATAAAAAATCTGGTCTGAAATGAAGAAATAAACAACTGACCTGA	840
Qy	991	TCTCT	GAGAAATATTACGTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA	1050
Db	841	TCTCT	GAGAAATATTACGTCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCA	900
Qy	1051	GTCACT	TGAAACTGTTTCCACATCAATAGTTGATGTCTGATACCAACGAATTCCTAAAAA	1110
Db	901	GTCACT	TGAAACTGTTTCCACATCAATAGTTGATGTCTGATACCAACGAATTCCTAAAAA	960
Qy	1111	GTGAC	CAGCTCTTAACAGCTAGCGAACGTAACTTAGACTTTACAGAGATTTATACGATCCCTC	1170
Db	961	GTGAC	CAGCTCTTAACAGCTAGCGAACGTAACTTAGACTTTACAGAGATTTATACGATCCCTC	1020
Qy	1171	GTGATA	AGCGTAAACTACTCTCAACAATCTCGATGCTTTTGGTATATATGAGCTATACCT	1230
Db	1021	GTGATA	AGCGTAAACTACTCTCAACAATCTCGATGCTTTTGGTATATATGAGCTATACCT	1080
Qy	1231	TAACT	TGGAAGAGTAGAGATATACAGTACACCAACCGTATCATAAACCGTTTATATGG	1290
Db	1081	TAACT	TGGAAGAGTAGAGATATACAGTACACCAACCGTATCATAAACCGTTTATATGG	1140
Qy	1291	GCAAC	GCGACCGGAAGGAGAGATGCTAGCTATCATTTAGCCGGTGGTGTGCTAGGCCGACG	1350
Db	1141	GCNAC	GCGACCGGAAGGAGAGATGCTAGCTATCATTTAGCTGGTGGTGCACGAGCGCAAC	1200
Qy	1351	AAAT	GGTTTC	1359
Db	1201	AGATT	GTAC	1209

RESULT 5

```

US-10-631-558-12
; Sequence 12, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sabni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

i OTHER INFORMATION: Hybrid cassette									
US-10-631-558-12									
Query Match 70.2%; Score 1165.8; DB 7; Length 2096;									
Best Local Similarity 99.4%; Pred. No. 0;									
Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;									
Qy	183	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGGTTGTT	242						
Db	588	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGGTTGTT	647						
Qy	243	AGCGTTCGTGTACTGTGTGAGGAGCGAATCAAGACATTAAGTCTTAATATTTTGAATC	302						
Db	648	AGCGTTCGTGTGTACTGTGTGAGGAGCGAATCAAGACATTAAGTCTTAATATTTTGAATC	707						
Qy	303	GATCTAAACATCAACGACCTGCTCATGTGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATCA	362						
Db	708	GATCTAAACATCAACGACCTGCTCATGTGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATCA	767						
Qy	363	AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTA	422						
Db	768	AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTA	827						
Qy	423	AAGGCTATTCAAGAACAAATTGATCGCTAAACGTTCACAGTAACGACGACTACTTTGAGGTC	482						
Db	828	AAGGCTATTCAAGAACAAATTGATCGCTAAACGTTCACAGTAACGACGACTACTTTGAGGTC	887						
Qy	483	ATTGATTTTGGCAAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC	542						
Db	888	ATTGATTTTGGCAAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC	947						
Qy	543	AAAGATGGTTTCGGTAAACCTTTGCCGACCAACCTGTCTCAAGAAATTTTGTCTAAGCGGACAT	602						
Db	948	AAAGATGGTTTCGGTAAACCTTTGCCGACCAACCTGTCTCAAGAAATTTTGTCTAAGCGGACAT	1007						
Qy	603	GTGGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG	662						
Db	1008	GTGGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG	1067						
Qy	663	GAATATACTGTACAGTTTACTCCCTTAAACCCGTATGACGATTTTCAGCACAGCTCTCAAA	722						
Db	1068	GAATATACTGTACAGTTTACTCCCTTAAACCCGTATGACGATTTTCAGCACAGCTCTCAAA	1127						
Qy	723	GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA	782						
Db	1128	GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA	1187						
Qy	783	GCTCAAGCACAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGCTGAC	842						
Db	1188	GCTCAAGCACAAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGCTGAC	1247						
Qy	843	TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACATGGATCAAGAG	902						
Db	1248	TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACATGGATCAAGAG	1307						
Qy	903	TTTACTTTACCGTGTTTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGCTCTG	962						
Db	1308	TTTACTTTACCGTGTTTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGCTCTG	1367						
Qy	963	AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG	1022						
Db	1368	AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG	1427						
Qy	1023	GAAAGCCGATGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCACCATCAAAATCGTT	1082						
Db	1428	GAAAGCCGATGATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCACCATCAAAATCGTT	1487						
Qy	1083	GATGTCGATACCAACGAATTCCTTAAAGGTGACGAGCTCTTACAGCTACGGAACGTAAC	1142						
Db	1488	GATGTCGATACCAACGAATTCCTTAAAGGTGAGGAGCTCTTACAGCTACGGAACGTAAC	1547						
Qy	1143	TTAGACTTTCAGAGATTTTATACGATCCCTCGGTGATAAGGCTAAACTACTCTACAACAATCTC	1202						

Db 1548 TTAGACTTCAGAGATTATATACGATCCTCGTGATAGGCTAAACTACTCTACAACAATCTC 1607
Qy 1203 GATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGCTAGAGGATAATACGATGAC 1262
Db 1608 GATGCTTTTGGTATTATGAGTATACCTTAACCTGGAAGCTAGAGGATAATACGATGAC 1667
Qy 1263 ACCAACCGTATCATACACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 1668 ACCAACCGTATCATACACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAC 1727
Qy 1323 CATTTAGCGGTGGTGTGAGCGCGAGCAAAATGGTTC 1359
Db 1728 CATTTAGCTGGTGGTGGCGGCAACAGATTGTAC 1764

RESULT 6

US-10-631-558-11
; Sequence 11, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-11

Query Match 69.2%; Score 1150.2; DB 7; Length 1782;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCAATTTGGTGT 242
Db 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCAATTTGGTGT 597
Qy 243 AGCGTTGCTGTACTGTGTAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 598 AGCGTTGCTGTACTGTGTAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGAATC 657
Qy 303 GATCTAAATCATGACCTGCTCATATGAGGAGAAAGACAGAGCAAGCTTTAAGTCAAAATCA 362
Db 658 GATCTAAATCATGACCTGCTCATATGAGGAGAAAGACAGAGCAAGCTTTAAGTCAAAATCA 717
Qy 363 AAACCAATTTGCTACTCATATGAGTGGCGGATGTCAATAAATTCAGAAAGCTGACTACTA 422
Db 718 AAACCAATTTGCTACTCATATGAGTGGCGGATGTCAATAAATTCAGAAAGCTGACTACTA 777
Qy 423 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAAACGAGCTACTTTGAGTTC 482
Db 778 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAAACGAGCTACTTTGAGTTC 837

RESULT 7

US-10-631-558-1
; Sequence 1, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha

Qy 483 ATTGATTTTGCAGAGATGCAACCATTTACTGATCGAAACGCGAAGGTCTACTTTGCTGAC 542
Db 838 ATTGATTTTGCAGAGATGCAACCATTTACTGATCGAAACGCGAAGGTCTACTTTGCTGAC 897
Qy 543 AAAGATGGTTCGGTAACTTTGCCGCAACCACTCTGCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 898 AAAGATGGTTCGGTAACTTTGCCGCAACCACTCTGCCAAGAAATTTTGTCTAAGCGGACAT 957
Qy 603 GTGGCGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 662
Db 958 GTGGCGCTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 1017
Qy 663 GAATATACTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722
Db 1018 GAATATACTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1077
Qy 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTAATA 782
Db 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTAATA 1137
Qy 783 GCTCAAGCAACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
Db 1138 GCTCAAGCAACAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
Qy 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
Db 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
Qy 903 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
Db 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1377
Qy 1023 GAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTTCAACATCAATACGTT 1082
Db 1378 GAAAGCCGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTTCAACATCAATACGTT 1437
Qy 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGGAGCTCTTAAACAGCTAGCGAAACGTAAC 1142
Db 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGGAGCTCTTAAACAGCTAGCGAAACGTAAC 1497
Qy 1143 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGGCTAACTACTCTACAAATCTC 1202
Db 1498 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAGGCTAACTACTCTACAAATCTC 1557
Qy 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCACGATGAC 1262
Db 1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATAATCACGATGAC 1617
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1677
Qy 1323 CATTTAGCGGTGGT 1337
Db 1678 CATTTAGCTTATGAT 1692

APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-10-631-558-1

Query Match : 69.2%; Score 1148.6; DB 7; Length 1245;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACACGCCAATGGTGT 242
Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACACGCCAATGGTGT 60
QY 243 ACGTTGCTGCTACTCTGTGAGGGGAGCAATCAACACATTAAGTCTTAATTTTGAATC 302
Db 61 ACGTTGCTGCTACTCTGTGAGGGGAGCAATCAACACATTAAGTCTTAATTTTGAATC 120
QY 303 GATCTAACTACGACCTGCTCATGTGAGGAGAAACACAGACGAAGCTTTAAGTCCAAATCA 362
Db 121 GATCTAACTACGACCTGCTCATGTGAGGAGAAACACAGACGAAGCTTTAAGTCCAAATCA 180
QY 363 AAACATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 422
Db 181 AAACATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAAACGACGACTTTGAGTGC 482
Db 241 AAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAAACGACGACTTTGAGTGC 300
QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGTGAGGAGAAACACAGACGAAGCTTTAAGTCCAAATCA 542
Db 301 ATTGATTTTGAAGCGATGCAACCATTTACTGTGAGGAGAAACACAGACGAAGCTTTAAGTCCAAATCA 360
QY 543 AAGATGTTTCCGTAACTTTGCGGACCCCAACCTGTCCAGAAATTTTGTGTAAGCGGACAT 602
Db 361 AAGATGTTTCCGTAACTTTGCGGACCCCAACCTGTCCAGAAATTTTGTGTAAGCGGACAT 420
QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY 663 GAATATAGTACAGTTTACTCCCTTAACCCCTGATGACGATTTTCAGACGAGTCTCAA 722
Db 481 GAATATAGTACAGTTTACTCCCTTAACCCCTGATGACGATTTTCAGACGAGTCTCAA 540
QY 723 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
Db 541 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
QY 783 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATAGATTTTATGACGCTGAC 842
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAAACCCAGGCTATAGATTTTATGACGCTGAC 660

QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
Db 661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
QY 903 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 721 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 963 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTAGCTCCTTAAAAAAGG 1022
Db 781 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTAGCTCCTTAAAAAAGG 840
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAAACCTGTTCAACATCAATACGTT 1082
Db 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAAACCTGTTCAACATCAATACGTT 900
QY 1083 GATGTCGATACCAAGCAATGCTTAAAGTAGAGGAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
Db 901 GATGTCGATACCAAGCAATGCTTAAAGTAGAGGAGCTCTTAAACAGCTAGCGAAGCTAAC 960
QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTTAACTACTCTACAACTCTC 1202
Db 961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTTAACTACTCTACAACTCTC 1020
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGAGAAAGTAGAGGATAATCAGATGAC 1262
Db 1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGAGAAAGTAGAGGATAATCAGATGAC 1080
QY 1263 ACCAACCGTATCATAAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGATGCTAGCTAT 1322
Db 1081 ACCAACCGTATCATAAACCGTTTATATGGGCAAGGACCCGGAAGGAGAGATGCTAGCTAT 1140
QY 1323 CATTTAGCCGGTGGT 1337
Db 1141 CATTTAGCCGTATGAT 1155
RESULT 8
US-10-631-558-3
; Sequence 3, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(777)
US-10-631-558-3

Query Match	19.0%	Score 316.4	DB 7	Length 777
Best Local Similarity	99.7%	Pred. No. 6.5e-89		
Matches 317	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy 1341	CAGCGCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTCACTCAAAAGCAAGCCCGGT	1400		
Db 1	CAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTCACTCAAAAGCAAGCCCGGT	60		
Qy 1401	TGTTTATGACAAATGGAACAACATATCAGATAAATCAAAGTGGGAGCGGACCTACTTAGGT	1460		
Db 61	TGTTTATGACAAATGGAACAACATATCAGATAAATCAAAGTGGGAGCGGACCTACTTAGGT	120		
Qy 1461	AATGTGTTGGTTTGTACTTGTATTGAGGAGACCGAGGTTTAACTGCGAAAGTAAACCT	1520		
Db 121	AATGTGTTGGTTTGTACTTGTATTGAGGAGACCGAGGTTTAACTGCGAAAGTAAACCT	180		
Qy 1521	GAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGCACCT	1580		
Db 181	GAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGCACCT	240		
Qy 1581	TATCAGCGCTCCTAAAGACTCCATCATCTGGGACGTGTACCTGCATCGGGGCTGGCGGAGGG	1640		
Db 241	TATCAGCGCTCCTAAAGACTCCATCATCTGGGACGTGTACCTGCATCGGGGCTGGCGGAGGG	300		
Qy 1641	AGAATAAGCTGTACCATC	1658		
Db 301	AGAATAAGCTGTACCATC	318		

```

RESULT 9
US-10-821-234-693/c
; Sequence 693, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821.234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: dt_seq_genes Version 1.0
; SEQ ID NO 693
; LENGTH: 8232
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-693

```

Query Match	18.8%	Score 312;	DB 7;	Length 8232;
Best Local Similarity	98.4%;	Pred. No. 6.7e-87;		
Matches 315; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;
Qy	1339	GTCAAGCGCAGCAAAATGTTTCAGCGCCAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCGCG	1398	
Db	7874	GGCAGGCTCAGCAAAATGTTTCAGCGCCAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCGCG	7815	
Qy	1399	GTGTGTTATGCAATGGAAACATCATCAGATATAATCAACAGTGGGAGCGGACCTACCTAG	1458	
Db	7814	GTGTGTTATGCAATGGAAACATCATCAGATATAATCAACAGTGGGAGCGGACCTACCTAG	7755	
Qy	1459	GTAATGTGTGTGTTTGTACTTGTATTGAGGGAAGCCGAGGTTTTTAACTGCGAAAGTAAAC	1518	
Db	7754	GCAATGCTGTGTTTGTACTTGTATTGAGGGAAGCCGAGGTTTTTAACTGCGAGAGTAAAC	7695	
Qy	1519	CTGAAGCTGAAGAGACTGTCTTTGACAAGTACACTTACCGAGTGGGTGACA	1578	
Db	7694	CTGAAGCTGAAGAGACTGTCTTTGACAAGTACACTTACCGAGTGGGTGACA	7635	
Qy	1579	CTTATGAGCGTCTCAAAGACTCATGATCATCTGGGACTGTACTCTCGATCGGGCTGGCGGAG	1638	

```

Db      7634 CTTATGAGCGTCTAAAGACTCCCATGATCTGGGA CTGTAC TCCTGC ATCGGG GCTTGG CGCAG 7575
QY      1639 GGAGAATAAGCTGTACCATC 1658
Db      7574 GGAGAATAAGCTGTACCATC 7555

RESULT 10
US-10-995-561-114
; Sequence 114, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-114
```

Query Match	18.8%	Score 311.6;	DB 7;	Length 2443;
Best Local Similarity	99.1%	Pred. No. 4.2e-87;		
Matches 314;	Conservative 1;	Mismatches 5;	Indels 0;	Gaps 0;
QY	1339	GTCAAGCGCAGCAAAATGGTTTCAGCCCGACGTC	CCCGTGGCGTGTCAAGCAAGCCCG	1398
Db	464	GGCAGGCTCAGCAAAATGGTTTCAGCCCGACGTC	CCCGTGGCGTGTCAAGCAAGCCCG	523
QY	1399	GTTCGTTATGCAATGGAAAAACATATCAGATAAAT	CAACAGTGGGAGCGACCTACCTAG	1458
Db	524	GTTCGTTATGCAATGGAAAAACATATCAGATAAAT	CAACAGTGGGAGCGACCTACCTAG	583
QY	1459	GTAATGTGTTGGTTGTACTTTGATGGAGAACG	CGAGGTTTTTAATCTGCGAGTAAAC	1518
Db	584	GCAATGCTTGGTTTGTACTTTGATGGAGAACG	CGAGGTTTTTAATCTGCGAGTAAAC	643
QY	1519	CTGAAGCTGAAGAGACTTGCTTTTGACAAAGTAC	ACTGGACACTTACCGNGTGGGTGACA	1578
Db	644	CTGAAGCTGAAGAGACTTGCTTTTGACAAAGTAC	ACTGGACACTTACCGAGTGGGTGACA	703
QY	1579	CTTATGAGCGTCTTAAAGACTCCATGATCTGGAC	TGTACTCTGCATCGGGGCTGGGCGAG	1638
Db	704	CTTATGAGCGTCTTAAAGACTCCATGATCTGGAC	TGTACTCTGCATCGGGGCTGGGCGAG	763
QY	1639	GGAGATAAGCTGTACCATC	1658	
Db	764	GGAGATAAGCTGTACCATC	783	

```

RESULT 11
US-10-995-561-105
; Sequence 105, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 2488

```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-105

Query Match      18.8%; Score 311.6; DB 7; Length 2488;
Best Local Similarity 98.1%; Pred. No. 4.3e-87;
Matches 314; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGGAGCAAAATGGTTTCAGCCCGAGTCCCGGTTGGTGTGTCAGTCAAGCAAGCCCG 1398
Db 464 GCGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTTGGTGTGTCAGTCAAGCAAGCCCG 523
QY 1399 GTTGTATGACATGGAAGAACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 524 GTTGTATGACATGGAAGAACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 583
QY 1459 GTAATGTGTTGTTTCTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 1518
Db 584 GCAATGCGTTGTTGTTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 643
QY 1519 CTGAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTTGGGAACTTACCGAGTGGTGACA 1578
Db 644 CTGAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTTGGGAACTTACCGAGTGGTGACA 1638
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 764 GGAGATAAGCTGTACCATC 783

RESULT 12
US-10-995-561-112
; Sequence 112, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 6510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-112

Query Match      18.8%; Score 311.6; DB 7; Length 6510;
Best Local Similarity 98.1%; Pred. No. 7.8e-87;
Matches 314; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGGAGCAAAATGGTTTCAGCCCGAGTCCCGGTTGGTGTGTCAGTCAAGCAAGCCCG 1398
Db 464 GCGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTTGGTGTGTCAGTCAAGCAAGCCCG 523
QY 1399 GTTGTATGACATGGAAGAACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 524 GTTGTATGACATGGAAGAACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 583
QY 1459 GTAATGTGTTGTTTCTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 1518
Db 584 GCAATGCGTTGTTGTTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 643
QY 1519 CTGAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTTGGGAACTTACCGAGTGGTGACA 1578
Db 644 CTGAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTTGGGAACTTACCGAGTGGTGACA 1638

RESULT 13
US-10-995-561-117
; Sequence 117, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117
; LENGTH: 7823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-117

Query Match      18.8%; Score 311.6; DB 7; Length 7823;
Best Local Similarity 98.1%; Pred. No. 8.7e-87;
Matches 314; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTGAGCGGAGCAAAATGGTTTCAGCCCGAGTCCCGGTTGGTGTGTCAGTCAAGCAAGCCCG 1398
Db 464 GCGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTTGGTGTGTCAGTCAAGCAAGCCCG 523
QY 1399 GTTGTATGACATGGAAGAACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458
Db 524 GTTGTATGACATGGAAGAACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 583
QY 1459 GTAATGTGTTGTTTCTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 1518
Db 584 GCAATGCGTTGTTGTTACTTGTATGGAGGAGCCGAGGTTTAACTCGGAAGTAAAC 643
QY 1519 CTGAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTTGGGAACTTACCGAGTGGTGACA 1578
Db 644 CTGAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTTGGGAACTTACCGAGTGGTGACA 1638
QY 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTGGGACTGTGGGACTGTGGGCGAG 1638
Db 704 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTGGGACTGTGGGCGAG 763
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 764 GGAGATAAGCTGTACCATC 783

RESULT 14
US-10-995-561-111
; Sequence 111, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
```

; LENGTH: 7848

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-995-561-111

Query Match 18.8%; Score 311.6; DB 7; Length 7848;

Best Local Similarity 98.1%; Pred. No. 8.7e-87;

Matches 314; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG 1398

Db 464 GGCAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG 523

Qy 1399 GTTGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458

Db 524 GTTGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 583

Qy 1459 GTAATGTGTGGTTTGTACTTGTATTGAGGAGCGGAGGTTTTAACTGCGAAAGTAAAC 1518

Db 584 GCAATGCGTTGGTTTGTACTTGTATTGAGGAGCGGAGGTTTTAACTGCGAGATAAC 643

Qy 1519 CTGAAGCTGAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578

Db 644 CTGAAGCTGAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 703

Qy 1579 CTTATGAGCGTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGGCTGGGCGAG 1638

Db 704 CTTATGAGCGTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGGCTGGGCGAG 763

Qy 1639 GGAGATAAGCTGTACCATC 1658

Db 764 GGAGATAAGCTGTACCATC 783

RESULT 15

US-10-995-561-113

; Sequence 113, Application US/10995561

; Publication No. US200502720541

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: CL001559

; CURRENT APPLICATION NUMBER: US/10/995,561

; CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 113

; LENGTH: 7935

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-995-561-113

Query Match 18.8%; Score 311.6; DB 7; Length 7935;

Best Local Similarity 98.1%; Pred. No. 8.8e-87;

Matches 314; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG 1398

Db 464 GGCAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG 523

Qy 1399 GTTGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 1458

Db 524 GTTGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAG 583

Qy 1459 GTAATGTGTGGTTTGTACTTGTATTGAGGAGCGGAGGTTTTAACTGCGAAAGTAAAC 1518

Db 584 GCAATGCGTTGGTTTGTACTTGTATTGAGGAGCGGAGGTTTTAACTGCGAGATAAC 643

Qy 1519 CTGAAGCTGAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578

Db 644 CTGAAGCTGAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 703

Qy 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCAATCGGGCTGGGCGAG 1638

Db 704 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCAATCGGGCTGGGCGAG 763

Qy 1639 GGAGATAAGCTGTACCATC 1658

Db 764 GGAGATAAGCTGTACCATC 783

Search completed: February 1, 2006, 14:43:13

Job time : 308.097 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 6825.35 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 1661
Sequence: 1 gcaaccgccagctagcc.....gaataagctgtaccatctaa 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_est4:*
5:	gb_est5:*
6:	gb_est6:*
7:	gb_est7:*
8:	gb_est8:*
9:	gb_gss1:*
10:	gb_gss2:*
11:	gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	312	18.8	480	2	BF993488 QVO-GN021
2	312	18.8	597	7	CN332340 170005325
3	312	18.8	657	7	CN332343 170005314
4	312	18.8	660	7	CN332334 170004554
5	312	18.8	666	7	CN332338 170006001
6	312	18.8	673	7	CN419479 170004706
7	312	18.8	676	1	AL603362 DKF2p686C
8	312	18.8	684	7	CN332332 170005326
9	312	18.8	693	8	DR004705 TC124975
10	312	18.8	706	7	CN332335 170005322
11	312	18.8	716	7	CN332342 170005339
12	312	18.8	751	1	AL706288 DKF2p686P
13	312	18.8	765	5	AX473413 DKF2p686I
14	312	18.8	911	5	EX398838 BX398838
15	312	18.8	7434	11	DQ039102 Homo sapi
16	312	18.8	7868	4	HSM806653
17	312	18.8	8411	4	HSM806992
18	311.6	18.8	1038	5	AX439175 Homo sapi
19	311	18.7	689	5	AX473425 DKF2p686L
20	311	18.7	719	5	AX473443 DKF2p686P
21	311	18.7	749	5	AX473395 DKF2p686D
22	311	18.7	759	5	AX473392 DKF2p686C

c	23	310.4	18.7	487	2	BF993485	BF993485	QVO-GN021
	24	310.4	18.7	494	3	BQ366579	BQ366579	QVO-GN021
	25	310.4	18.7	683	5	BX501294	BX501294	DKF2p729A
	26	310.4	18.7	693	1	AL706197	AL706197	DKF2p686I
	27	310.4	18.7	699	1	AL706221	AL706221	DKF2p686J
	28	310.4	18.7	748	5	BX473391	BX473391	DKF2p686C
	29	310.4	18.7	756	5	BX473407	BX473407	DKF2p686G
	30	310.4	18.7	913	5	BX380583	BX380583	BX380583
	31	310.4	18.7	1026	5	BX386270	BX386270	BX386270
	32	310.4	18.7	1050	5	BX417945	BX417945	BX417945
	33	310.4	18.7	7777	4	CR749316	CR749316	Homo sapi
	34	310.4	18.7	7885	4	CR749281	CR749281	Homo sapi
	35	310	18.7	676	5	BX510005	BX510005	DKF2p686I
	36	310	18.7	708	8	DR005445	DR005445	TC109245
	37	310	18.7	7434	11	DQ039103	DQ039103	Pan trogl
	38	308.8	18.6	378	2	BF751799	BF751799	RC3-BN042
	39	308.8	18.6	427	2	BF988908	BF988908	IL5-GN017
	40	308.8	18.6	8121	4	CR749317	CR749317	Homo sapi
	41	307.8	18.5	712	7	CR791105	CR791105	DKF2p469C
	42	304.4	18.3	321	1	AW068245	AW068245	cn23g07.Y
	43	303.6	18.3	925	5	BX380854	BX380854	BX380854
	44	303	18.2	670	5	BX473387	BX473387	DKF2p686B
	45	301	18.1	684	7	CN332337	CN332337	170005326

ALIGNMENTS

RESULT 1
BF993488
LOCUS BF993488 480 bp mRNA linear EST 23-JAN-2001
DEFINITION QVO-GN0216-011100-466-g03 GN0216 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF993488
VERSION BF993488.1 GI:12399811
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QVO&t=2-QVO-GN0216-011100-466-g03&t3=2000-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 479.
Location/Qualifiers
1. 480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0216"

/note="Organ: placenta normal; Vector: puc18; Site:1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 2; Length 480;
Best Local Similarity 98.4%; Pred. No. 3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTCCAGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCAAGCCCG 1398
DB 124 GCGAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCAAGCCCG 183

QY 1399 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
DB 184 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 243

QY 1459 GTAATGTTGGTTTGTACTTGTATGGAGGAGCGGAGTTTAACTCGGAAGTAAC 1518
DB 244 GCAATGCGTGTGTGTACTTGTATGGAGGAGCGGAGTTTAACTCGGAGTAAC 303

QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578
DB 304 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 363

QY 1579 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 1638
DB 364 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 423

QY 1639 GGAGATAAGCTGTACCATC 1658
DB 424 GGAGATAAGCTGTACCATC 443

RESULT 2

CN332340

LOCUS CN332340 597 bp mRNA linear EST 16-MAY-2004

DEFINITION 1700532545117 GRN_EB Homo sapiens CDNA 5', mRNA sequence.

ACCESSION CN332340

VERSION CN332340.1 GI:47332274

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 597)
AUTHORS Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
TITLE Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

PUBMED 15146197

COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 597 Std Error: 0.00.

FEATURES

source

1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"

/clone lib="GRN_EB"
/note="Oligo dt primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 7; Length 597;
Best Local Similarity 98.4%; Pred. No. 3.2e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTCCAGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCAAGCCCG 1398
DB 255 GCGAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCAAGCCCG 314

QY 1399 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
DB 315 GTTGTATGCAATGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 374

QY 1459 GTAATGTTGGTTTGTACTTGTATGGAGGAGCGGAGTTTAACTCGGAAGTAAC 1518
DB 375 GCAATGCGTGTGTGTACTTGTATGGAGGAGCGGAGTTTAACTCGGAGTAAC 434

QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 1578
DB 435 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 494

QY 1579 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 1638
DB 495 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAG 554

QY 1639 GGAGATAAGCTGTACCATC 1658
DB 555 GGAGATAAGCTGTACCATC 574

RESULT 3

CN332343

LOCUS CN332343 657 bp mRNA linear EST 16-MAY-2004

DEFINITION 17005331483454 GRN_EB Homo sapiens CDNA 5', mRNA sequence.

ACCESSION CN332343

VERSION CN332343.1 GI:47332277

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 657)
AUTHORS Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
TITLE Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

PUBMED 15146197

COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 657 Std Error: 0.00.

FEATURES

source

1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone lib="GRN_EB"
/note="Oligo dt primed, full-length enriched cDNA library

from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions.²⁸

ORIGIN	Query Match	18.8%	Score 312;	DB 7;	Length 657;
	Best Local Similarity	98.4%	Pred. No. 3.3e-78;		
	Matches 315;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
Qy	1339	GTCAAGGCGCAGCAAAATGGTTTCAGGCCCCAGTCCCGGTGGGTGTGTCAGTCAAAGCAAGGCCCG	1398		
Db	287	GGCAGGCTCAGCAAAATGGTTTCAGGCCCCAGTCCCGGTGGGTGTGTCAGTCAAAGCAAGGCCCG	346		
Qy	1399	GTTCGTTATGCAATCGGAAACATCTATCAGATAAATCAACAGTGGGAGCGGACCTTACCTAG	1458		
Db	347	GTTCGTTATGCAATCGGAAACATCTATCAGATAAATCAACAGTGGGAGCGGACCTTACCTAG	406		
Qy	1459	GTAATGTGTTGGTTTGTACTTTGTTATGGAGGAAGCCGAGGTTTTTAACCTGCGAAAGTAAAC	1518		
Db	407	GCAATGCGTTGGTTTGTACTTTGTTATGGAGGAAGCCGAGGTTTTTAACCTGCGAGAGTAAAC	466		
Qy	1519	CTGAAGCTGAAGAGACCTTGCTTTTGCAAGTACACTGGGAAACACTTACCGAGTGGGTGACA	1578		
Db	467	CTGAAGCTGAAGAGACCTTGCTTTTGCAAGTACACTGGGAAACACTTACCGAGTGGGTGACA	526		
Qy	1579	CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGTCATCGGGGCTGGGCGAG	1638		
Db	527	CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGTCATCGGGGCTGGGCGAG	586		
Qy	1639	GGAGAAATAGCTGTACCATC	1658		
Db	587	GGAGAAATAGCTGTACCATC	606		

RESULT 4	CN332334	660 bp	linear	EST 16-MAY-2004
LOCUS	CN332334	17000455431365	GRN_EB Homo sapiens	cDNA 5', mRNA sequence.
DEFINITION	CN332334			
ACCESSION	CN332334			
VERSION	CN332334.1	GI:47332268		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 660)			
	Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M. S., Mandalem, R., Lebkowsk, J and Stanton, L. W.			
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation			
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)			
PUBLISHED	15146197			
COMMENT	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 660 Std Error: 0.00.			

```

FEATURES
source
insert length: 660      stat error: 0.00.
Location/Qualifiers
1. .660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone.lib="GRN_EB"
/note="Oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free

```

conditions."

Query Match	18.8%	Score 312;	DB 7;	Length 660;
Best Local Similarity	98.4%;	Pred. NO. 3.3e-78;		
Matches 315;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
Qy	1339	GTCCAGCGCAGCAAAATGGTTTCAGCCCGCCAGTCCCGGGTGCCTGCAGTCAAGCAAGCCCG	1398	
Db	325	GGCAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGGTGCCTGCAGTCAAGCAAGCCCG	384	
Qy	1399	GTTCGTTATGACAATGGAAAAACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG	1458	
Db	385	GTTCGTTATGACAATGGAAAAACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG	444	
Qy	1459	GTAATGTGTGGTTTGTACTGTGTATGGAGAAAGCCGAGGTTTTTAACCTGCGAAGTAAAC	1518	
Db	445	GCAATGCGTTGGTTTGTACTGTGTATGGAGAAAGCCGAGGTTTTTAACCTGCGAGTAAAC	504	
Qy	1519	CTCAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACATTTACCAGTGGGTGACA	1578	
Db	505	CTCAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACATTTACCAGTGGGTGACA	564	
Qy	1579	CTTTATGAGCGCTCCTAAAGACTCCATGATCTGGGACCTGTACTCTGCATCGGGGCTGGGCGAG	1638	
Db	565	CTTTATGAGCGCTCCTAAAGACTCCATGATCTGGGACCTGTACTCTGCATCGGGGCTGGGCGAG	624	
Qy	1639	GGAGAAATAAGCTGTACCATC	1658	
Db	625	GGAGAAATAAGCTGTACCATC	644	

RESULT 5	CN332338	666 bp	linear	EST 16-MAY-2004
CN332338	17000600181351	GRN_PRENU	Homo sapiens	cdna 5', mRNA sequence.
LOCUS	CN332338			
DEFINITION	CN332338			
ACCESSION	CN332338			
VERSION	CN332338.1	GI:47332272		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Hominidae; Homo.			
REFERENCE	1 (bases 1 to 666)			
AUTHORS	Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J and Scanton, L. W.			
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation			
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)			
PUBMED	15146197			
COMMENT	Contact: Brandenberger R			
	Regenerative Medicine			
	Genon Corporation			
	230 Constitution Drive, Menlo Park, CA 94025, USA			
	Tel: 650 473 8658			
	Fax: 650 473 7760			
	Email: rbrandenberger@genon.com			
	Insert Length: 666			Std Error: 0.00.

```

FEATURES
source
insert length: 666 Std. Error: 0.00.
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN PRENEU"
/note="oligo dt' primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic and mitogens."
ORIGIN

```

```
Query Match      18.8%; Score 312; DB 7; Length 666;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTCAGCGCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCG 1398
Db 297 GCGAGGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCG 356
QY 1339 GTTGTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 357 GTTGTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 416
QY 1459 GTAATGCTGTGTTGTTACTTGTATGGAGGAGCCGAGTGTAACTCGGAAAGTAAC 1518
Db 417 GCAATGCTGTGTTGTTACTTGTATGGAGGAGCCGAGTGTAACTCGGAGGTAAC 476
QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 1578
Db 477 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 536
QY 1579 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGGGAG 1638
Db 537 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGGGAG 596
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 597 GGAGATAAGCTGTACCATC 616

RESULT 6
CN419479      673 bp mRNA linear EST 16-MAY-2004
LOCUS        17000470666007 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION   CN419479
ACCESSION    CN419479
VERSION      CN419479.1 GI:47407073
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE    1 (bases 1 to 673)
AUTHORS      Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE        Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL      Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED       15146197
COMMENT      Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 673 Std Error: 0.00.
FEATURES     Location/Qualifiers
source       1..673
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cells, embryoid bodies
            derived from H1, H7 and H9 cells"
            /clone_lib="GRN EB"
            /note="oligo dt primed, full-length enriched cDNA library
            from embryoid body outgrowths derived from hES cell lines
            H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
            conditions."

ORIGIN
Query Match      18.8%; Score 312; DB 7; Length 673;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;

QY 1339 GTCAGCGCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCG 1398
Db 193 GGCAGGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCG 252
QY 1399 GTTGTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 253 GTTGTATGACAAATGGAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 312
QY 1459 GTAATGCTGTGTTGTTACTTGTATGGAGGAGCCGAGTGTAACTCGGAAAGTAAC 1518
Db 313 GCAATGCTGTGTTGTTACTTGTATGGAGGAGCCGAGTGTAACTCGGAGGTAAC 372
QY 1519 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 1578
Db 373 CTGAAGCTGAAGAGACTTCTTTGACAAGTACACTCGGGAACACTTACCGAGTGGGTGACA 432
QY 1579 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGGGAG 1638
Db 433 CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGGGAG 492
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 493 GGAGATAAGCTGTACCATC 512

RESULT 7
AL603362      676 bp mRNA linear EST 04-SEP-2003
LOCUS        DKFZp686C197_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION   DKFZp686C197_5', mRNA sequence.
ACCESSION    AL603362
VERSION      AL603362.1 GI:15166868
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE    1 (bases 1 to 676)
AUTHORS      Ansorge, W., Wirkner, U., Mewes, W., Weil, B. and Wiemann, S.
TITLE        EST (Ansorge, W., Wirkner, U., Mewes, H.W., Weil, B. and Wiemann, S.)
JOURNAL      Unpublished (1999)
COMMENT      Contact: MIPS
MIPS
Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp686C197) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES     Location/Qualifiers
source       1..676
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp686C197"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="686 (synonym: hlcc3)"
            /note="Vector: pTriplex2; Site_1: SfIIA; Site_2: SfiIB;
            cDNA-collection"

ORIGIN
Query Match      18.8%; Score 312; DB 1; Length 676;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
```

Qy	1339	GT	CAGGCGCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG	1398
Db	312	GG	CAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAAGCAAGCCCG	371
Qy	1399	GT	TGTTATGCAATGGAAACATCTACAGATAAATCAACAGTGGAGCGGACCTACCTAG	1458
Db	372	GT	TGTTATGCAATGGAAACATCTACAGATAAATCAACAGTGGAGCGGACCTACCTAG	431
Qy	1459	GT	AATGTTGGTTTGTACTTGTATCGAGGAGCGGAGTGTTTAACTCGGAAGTAAAC	1518
Db	432	GCA	ATGGTGGTTTGTACTTGTATCGAGGAGCGGAGTGTTTAACTCGGAGAGTAAC	491
Qy	1519	CT	GAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA	1578
Db	492	CT	GAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA	551
Qy	1579	CT	TATGAGCGTCCCTAAAGATCCATGATCTGGGACTGTGTACTCTGCATCGGGGCTGGGCGAG	1638
Db	552	CT	TATGAGCGTCCCTAAAGATCCATGATCTGGGACTGTGTACTCTGCATCGGGGCTGGGCGAG	611
Qy	1639	GG	AGAAATAGCTGTACCATC	1658
Db	612	GG	AGAAATAGCTGTACCATC	631

DR004705	693 bp	mRNA	linear	EST 17-MAY-2005
LOCUS				
DEFINITION	TC124975 Human placenta, large insert, pCMV expression library Homo sapiens cDNA clone TC124975 5', similar to Homo sapiens fibronectin 1 (FN1), transcript variant 3, mRNA sequence.			
ACCESSION	DR004705			
VERSION	DR004705.1	GI:62624578		
KEYWORDS	EST.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 693)			
AUTHORS	Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L., Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M., Zhang,X., Jay,G. and He,W.			
TITLE	High-throughput cloning of full-length human cDNAs directly from cDNA libraries optimized for large and rare transcripts			
JOURNAL	Unpublished (2005)			
COMMENT	Contact: Kovacs, KF			

Please contact OriGene for access.
OriGene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
<http://www.origene.com>
Seq primer: pCW6 5prime forward vector primer, OriGene
Technologies Inc.
Location/Qualifiers
source 1. .693

```

/db_xref="taxon:9606"
/clone="TCL24975"
/tissue type="Placenta"
/clone_lib="Human placenta, large insert, pCMV expression
library"
/notes="vector: pCMV6-XL4; Site 1: EcoRI; Site 2: XhoI/Sall"

```

compatible end ligation; Oligo-dT primed reverse transcription optimized for large and GC rich mRNA transcripts, cDNA size selection, optimized ligation for large inserts into mammalian expression vector, random clones selected for end sequence verification of full-length genes"

/note="oligo dT primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

```

ORIGIN
Query Match      18.8%; Score 312; DB 8; Length 693;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1339 GTCAGCGCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGCTGTGTCAGTCAAGCAAGCCCG 1398
Db 323 GCGAGGCTCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGCTGTGTCAGTCAAGCAAGCCCG 382
QY 1399 GTTGTTATGCAATGGAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 1458
Db 383 GTTGTTATGCAATGGAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 442
QY 1459 GTAATGTTGTTGGTTTGTACTTTATGGAGGAGCCGAGGTTTAACTGCGAAGTAAC 1518
Db 443 GCATCGGTTGGTTTGTACTTTATGGAGGAGCCGAGGTTTAACTGCGAAGTAAC 502
QY 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAACACTTACCGAGTGGTGACA 1578
Db 503 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAACACTTACCGAGTGGTGACA 562
QY 1579 CTTATGAGGCTCTTAAGACTCCATGATCTGGGACTGTACTGTCATCGGGCTGGGCGAG 1638
Db 563 CTTATGAGGCTCTTAAGACTCCATGATCTGGGACTGTACTGTCATCGGGCTGGGCGAG 622
QY 1639 GGAGATAAGCTGTACCATC 1658
Db 623 GGAGATAAGCTGTACCATC 642

```

```

RESULT 10
CN32335
LOCUS 706 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532295426 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN32335
VERSION CN32335.1 GI:47332269
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 706)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 706 Std Error: 0.00.
Location/Qualifiers
1..706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
FEATURES
source

```

```

RESULT 11
CN32342
LOCUS 716 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000533983768 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN32342
VERSION CN32342.1 GI:47332276
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 716)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 716 Std Error: 0.00.
Location/Qualifiers
1..716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/notes="oligo dT primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines

```

```

FEATURES
source

```

H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Query Match 18.8%; Score 312; DB 7; Length 716;
Best Local Similarity 98.4%; Pred. No. 3.3e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCCG 1398
Db 235 GGCAGGCTCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCCG 294

Qy 1399 GTTGTATACATGCAATGCAAAACACTATCAGATAAATCAACAGTGGGCGGACCTACCTAG 1458
Db 295 GTTGTATGACAATGCAAAACACTATCAGATAAATCAACAGTGGGCGGACCTACCTAG 354

Qy 1459 GTAATGTGTGTTTGTACTTGTATGAGGAGCCGAGGCTTTAACTGCGAAGTAAC 1518
Db 355 GCAATGCGTGTGTTGTACTTGTATGAGGAGCCGAGGCTTTAACTGCGAAGTAAC 414

Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGTGACA 1578
Db 415 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGTGACA 474

Qy 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGATCGGGCTGGGCGAG 1638
Db 475 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGATCGGGCTGGGCGAG 534

Qy 1639 GGAGAAATAGCTGTACCATC 1658
Db 535 GGAGAAATAGCTGTACCATC 554

RESULT 12

AL706288

LOCUS

DEFINITION

AL706288

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the CDNA sequencing
consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp686P033) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES

source

1. .751
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686P033"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;

ORIGIN

CDNA-collection"

Query Match 18.8%; Score 312; DB 1; Length 751;
Best Local Similarity 98.4%; Pred. No. 3.4e-78;
Matches 315; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1339 GTCAGGCGCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCCG 1398
Db 359 GGCAGGCTCAGCAAAATGGTTTCAGCCCCCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCCG 418

Qy 1399 GTTGTATGCAATGCAAAACACTATCAGATAAATCAACAGTGGGCGGACCTACCTAG 1458
Db 419 GTTGTATGCAATGCAAAACACTATCAGATAAATCAACAGTGGGCGGACCTACCTAG 478

Qy 1459 GTAATGTGTGTTTGTACTTGTATGAGGAGCCGAGGCTTTAACTGCGAAGTAAC 1518
Db 479 GCAATGCGTGTGTTGTACTTGTATGAGGAGCCGAGGCTTTAACTGCGAAGTAAC 538

Qy 1519 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGTGACA 1578
Db 539 CTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCGAGTGGTGACA 598

Qy 1579 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGATCGGGCTGGGCGAG 1638
Db 599 CTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGATCGGGCTGGGCGAG 658

Qy 1639 GGAGAAATAGCTGTACCATC 1658
Db 659 GGAGAAATAGCTGTACCATC 678

RESULT 13

BX473413

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the CDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp686I15162) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES

source

1. .765
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686I15162"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:47:34 ; Search time 115.543 Seconds
(without alignments)
12632.698 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaaccccgccagctagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delgap 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09940235/runat 27012006_144217_27548/app_query.fasta_1.7708
-DB=A Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CNG_1_1_846 @runat 27012006_144217_27548 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	413	ADL92189	Adl92189 Streptoki
2	1984	67.5	414	AAR10194	Aar10194 Streptoki
3	1984	67.5	414	AAR63120	Aar63120 Streptoki
4	1984	67.5	414	AAV24794	Aay24794 Streptoco
5	1984	67.5	414	AAW94664	Aaw94664 Streptoco
6	1984	67.5	414	AAV01556	Aay01556 Native st
7	1984	67.5	414	AAV90282	Aay90282 S. equisi
8	1984	67.5	795	AAV24797	Aay24797 Streptoki
9	1981	67.4	531	AAR11829	Aar11829 FB-FB-SK

10	1980	67.4	414	2	AAW94665	Aaw94665 Streptoco
11	1966.5	66.9	813	2	AAW21728	Aaw21728 Wild type
12	1966.5	66.9	1194	2	AAW21726	Aaw21726 Streptoki
13	1963	66.8	415	9	AE93425	Aeb93425 Streptoco
14	1958	66.6	414	2	AAW86143	Aaw86143 Streptoki
15	1958	66.6	414	3	AAW01295	Aab01295 Wild type
16	1958	66.6	440	2	AAR12889	Aar12889 Streptoki
17	1958	66.6	483	2	AAR12885	Aar12885 Factor Xa
18	1958	66.6	483	2	AAR12522	Aar12522 Factor Xa
19	1958	66.6	499	2	AAR12891	Aar12891 Streptoki
20	1958	66.6	859	2	AAR12893	Aar12893 OmpA str
21	1954	66.5	414	2	AAW20202	Aaw20202 S.equisim
22	1951.5	66.4	415	3	AAV50870	Aay50870 Streptoco
23	1951.5	66.4	415	3	AAV99593	Aay99593 Streptoco
24	1947	66.2	414	2	AAW86144	Aaw86144 De-immuni
25	1947	66.2	414	3	AAW01296	Aab01296 Altered s
26	1942	66.1	440	2	AAW06377	Aar06377 Streptoki
27	1936.5	65.9	1194	2	AAW21724	Aaw21724 Modified
28	1936.5	65.9	1194	2	AAW21725	Aaw21725 Modified
29	1928	65.6	413	2	AAV25020	Aay25020 Streptoki
30	1928	65.6	413	5	ABB80012	Abb80012 Streptoki
31	1928	65.6	413	6	ABG74199	Abg74199 Represent
32	1924	65.5	372	2	AAAR10200	Aar10200 Streptoki
33	1916	65.2	372	2	AAAR10197	Aar10197 Streptoki
34	1907.5	64.9	371	2	AAAR10195	Aar10195 Streptoki
35	1907	64.9	374	2	AAAR10198	Aar10198 Streptoki
36	1904	64.8	414	2	AAW74940	Aaw74940 Mutant st
37	1902	64.7	384	3	AAV84007	Aay84007 Amino aci
38	1900	64.6	800	2	AAW21723	Aaw21723 Plasminog
39	1900	64.6	1181	2	AAW21727	Aaw21727 Streptoki
40	1898	64.6	372	3	AAV84006	Aay84006 Amino aci
41	1887	64.2	401	3	AAV84004	Aay84004 Amino aci
42	1887	64.2	413	3	AAV84005	Aay84005 Amino aci
43	1879	63.9	747	2	AAAR12894	Aar12894 Met-core
44	1875	63.8	369	2	AAAR12892	Aar12892 Truncated
45	1855	63.1	440	1	AAAP50620	Aap50620 Sequence

ALIGNMENTS

RESULT 1

ADL92189

ID ADL92189 standard; protein; 413 AA.

AC ADL92189;

DT 20-MAY-2004 (first entry)

DE Streptokinase protein sequence.

KW harvesting; recombinant; host cell; N-terminal leader peptide;
KW pre-peptide; lantibiotic; post-translational modification;
KW pharmaceuticals; vaccine; immunogenic.

OS Unidentified.

PN WO2003099862-A1.

XX 04-DEC-2003.

PF 26-MAY-2003; 2003WO-NL000389.

PR 24-MAY-2002; 2002EP-00077060.

PR 07-FEB-2003; 2003US-00360101.

XX (NANO-) APPLIED NANOSYSTEMS BV.

PI Moll CN, Leenhouts CJ, Kuipers OP, Driessen AJM;

XX WPI; 2004-042770/04.

PT Harvesting a desired polypeptide produced by a recombinant host cell, for
producing pharmaceuticals, comprises selecting a recombinant nucleic acid

PT comprising nucleic acid fragments encoding a leader peptide and the
 XX polypeptide.

PS Claim 4; Page 82-83; 109pp; English.

XX The invention relates to a novel method for harvesting a (poly)peptide
 CC produced by a recombinant host cell. The novel method involves selecting
 CC a cell comprising a first nucleic acid encoding a leader peptide and a
 CC second nucleic acid fragment encoding the desired (poly)peptide. The
 CC first and second fragments are within the same open reading frame of the
 CC first nucleic acid and the leader peptide is functionally equivalent to
 CC an N-terminal leader peptide found with the pre-peptide of a lantibiotic.
 CC The host cells and nucleic acids are useful for producing, harvesting and
 CC post-translational modification of polypeptides. The polypeptides may be
 CC used in the production of pharmaceuticals, e.g. as antigen for vaccine or
 CC immunogenic composition. This sequence represents a polypeptide relating
 CC to the novel method of the invention.

XX SQ Sequence 413 AA;

Alignment Scores:

Pred. No.: 3,93e-178 Length: 413
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 8 Gaps: 0

US-09-940-235-10 (1-1661) x ADL92189 (1-413)

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCTGCTCCTCATCTGTCAACACAGCCAATTTGGTTGTT 242
 DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnSerGlnLeuVal 20
 QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGATCAAGACATTAGTCTTAAATTTTGGAAATC 302
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGCAGCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 362
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCATTTGCTGATGATGCGCGATGTCCATATAACTTGAGAAAGCTGACTTACTA 422
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
 QY 423 AAGGCTATTCAAGACAAATTGATCGCTAACCGTCCACAGTAACGACGACTACTTTGAGGTC 482
 DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
 QY 483 ATTGATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGCTTACTTTGCTGAC 542
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGTTGCGTACCTTGGCCGACCACTGCTCCAAAGATTTTGTCTAAGCGCACAT 602
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGGCGTTTAGACCATATAAGAAACCAATACAAACCAAGCGCAAAATCTGTTGATGTG 662
 DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATACGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTCAGACAGTCTCAAA 722
 DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACTAGCTATTCAACACTGCTATCGTGTGACCATCATCATCTCAAGAAATCTACTA 782
 DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
 QY 783 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGAGTATATGACGTGAC 842
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220

QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAGCCGATCATCCCTTTGATCGAGTCACCTGAAACTGTTCCACATCAAAATACGTT 1082
 DB 281 GluLysProTyrAspPropheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACGAATTTGTAAGAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 1142
 DB 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGATTTATACGATCTCTCGTGAATAAGCTTAACTACTCTTCAACAATCTC 1202
 DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTGCTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
 DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATAACCGTTTATATGGCAAGCGACCCCAAGAGAGAGATGCTAGCTAT 1322
 DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 DB 381 HisLeuAla 383
 RESULT 2
 AAR10194
 ID AAR10194 standard; protein; 414 AA.
 XX
 AC AAR10194;
 DT 28-MAR-1991 (first entry)
 XX
 DE Streptokinase encoded by synthetic gene.
 XX streptokinase; thrombolytic agent; myocardial infarction.
 OS Synthetic.
 XX
 PN EP407942-A.
 XX
 PD 16-JAN-1991.
 XX
 PF 11-JUL-1989; 89JP-00179432.
 XX
 PR 11-JUL-1989; 89JP-00179432.
 PR 27-NOV-1989; 89JP-00307957.
 PR 11-APR-1990; 90JP-00096830.
 XX
 PA (SAKA) OTSUKA PHARM FACTOR.
 XX
 PI Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
 XX
 DR WPI: 1991-016179/03.
 DR N-PSDB; AAQ10230.
 XX
 FT Synthetic gene encoding streptokinase - scale, high purity prodn. of
 FT streptokinase used as a thrombolytic agent.
 XX
 PS Claim 1; Page 51; 76pp; English.
 XX

CC Streptokinase and its derivatives can be produced in large quantities
CC with high purity for use as thrombolytic agents in patients with lung
CC thrombus or myocardial infarction. See also AAR10195-R10200
XX
SQ Sequence 414 AA;

Alignment Scores:
Pred. No.: 3,93e-178 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAR10194 (1-414)

QY	183	ATTGCTGGACCTGAGTGGCTGTAGACCTGCTCTGTCACACACGCCAATGGTGT	242
Db	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
QY	243	AGCGTTGCTGTACTGCTTGAGGGGACCAATCAAGACATTAGTCTTAAATTTTGAATC	302
Db	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluile	40
QY	303	GATCTAACATCAACGCTCTCATGAGGAGAAAGACAGACGAGCTTAACTCCAAATCA	362
Db	41	AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	60
QY	363	AAACATTGCTACTGATAGTGGCGGATGTCATATAAAGCTGAGAAAGCTGACTACTA	422
Db	61	LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu	80
QY	423	AAGGCTATTCAAGACCAATGATCGCTTAAGCTCCACAGTAACGACGACTACTTGGGTC	482
Db	81	LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspPheGluVal	100
QY	483	ATTGATTTTGAAGCGATGCAACATTACTGTATCGAAACGGCAAGGTCTACTTTGTGAC	542
Db	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyPheAlaAsp	120
QY	543	AAAGATGGTTCGGTAACCTTGGCGACCCAACTGTGTCAGAAATTTTGTAAACGGACAT	602
Db	121	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	140
QY	603	GTGCGGTTAGACCATATAAGAAACCAATCAAAACCAAGCAATCTGTTGATGTG	662
Db	141	ValArgValArgProTyLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	663	GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACCAGGTCTCAA	722
Db	161	GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
QY	723	GATACCTAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCTCAAGAATTACTA	782
Db	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	200
QY	783	GCTCAAGCAAAAGCATTTTAAACAAACCCAGGCTATAGCATTTATGACGTCGAC	842
Db	201	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGluArgAsp	220
QY	843	TCCTCAATCGTCATCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	902
Db	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	903	TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
Db	241	PheThrTyArgValLysAsnArgGluGlnAlaTyArgIleAsnLysLysSerGlyLeu	260
QY	963	AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAGGG	1022
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyTyValLeuLysLysGly	280
QY	1023	GAAAGCCGTATGATCCCTTTGATCCAGTCACTTGAACTGTTCCACCAATACGTT	1082

Db	281	GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal	300
QY	1083	GATCTCGATACCAACGAATTCTCTAAAGTGCAGCAGCTCTTAAACAGCTAGCGACGTAC	1142
Db	301	AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
QY	1143	TTAGACTTCAGAGATTTATATACGATCCTCGTATAGGCTAAACTACTCTACAACAATCTC	1202
Db	321	LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu	340
QY	1203	GATCCTTTTGTATATATGAGTACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC	1262
Db	341	AspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAsnHisAspAsp	360
QY	1263	ACCAACCGTATCAACCGTTTATATGCGCAAGCGACCGAGGAGAGAAATGCTAGCTAT	1322
Db	361	ThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSerTy	380
QY	1323	CATTAGCC 1331	
Db	381	HisLeuAla 383	
RESULT 3			
AAR63120			
ID	AAR63120 standard; protein; 414 AA.		
XX	AAR63120;		
XX	16-OCT-2003 (revised)		
DT	25-MAR-2003 (revised)		
DT	16-NOV-1994 (first entry)		
XX	Streptokinase.		
XX	Streptococcus SK; Streptococcus equisimilis; plasminogen;		
KW	myocardial infarction.		
XX	Streptococcus dysgalactiae subsp. equisimilis.		
OS	Streptococcus dysgalactiae subsp. equisimilis.		
FH	Key Location/Qualifiers		
FT	Region	1..352	
FT	/note= "claim 3, see CC"		
FT	Region	14..414	
FT	/note= "claim 1, see CC"		
FT	Region	120..352	
FT	/note= "claim 3, see CC"		
FT	Region	244..414	
FT	/note= "claim 3, see CC"		
FT	Region	244..352	
FT	/note= "claim 2, see CC"		
XX	WO9407992-A1.		
PN	14-APR-1994.		
PD	05-OCT-1993; 93WO-US009502.		
PF	05-OCT-1992; 92US-00956692.		
XX	29-SEP-1993; 93US-00128299.		
PR	(GEHO) GEN HOSPITAL CORP.		
XX	(HARD) HARVARD COLLEGE.		
PA	Reed GL;		
XX	WPI; 1994-135561/16.		
DR	DNA encoding a polypeptide which binds to plasminogen and corresponds to		
XX	region of streptokinase - useful to detect plasminogen in a sample and to		
PT	treat myocardial infarction.		
XX	Disclosure; Page 40-41; 62pp; English.		
PS			

```

XX Nucleic acid comprising a sequence encoding amino acids 14-414 of
CC streptokinase, which binds to plasminogen and does not have a sequence
CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 414 AA;
Alignment Scores:
Pred. No.: 3,93e-178 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-10 (1-1661) x AAR63120 (1-414)
QY 183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAATTTGGTGT 242
DB 1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 243 AGCGTTGCTGCTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAGACAGCAAGCGCTTAAGTCCAAATCA 362
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTGATGATGGCGGATGTCACATAAATCTGAGAAAGCTGACTACTA 422
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGGCTATTCAAGAACTGATCGCTAACCGTCCACAGTAACAGCACTACTTTGAGGTC 482
DB 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
QY 483 ATTGATTTGCAAGCGATCAACCTACTGATGCAAAACGCAAGCTCTACTTTGCTGAC 542
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAGATGCTTTCGTTAACTTGGCGGACCCACCTGCTCCAGAAATTTTGTAGCGGACAT 602
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 662
DB 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTACAGCAGGTCTCAAA 722
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATACTAGCTATTGAAAACACTACTATCGTGTGACCATCACATCTCAAGAAATTACTA 782
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTAACTGAC 842
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleThrGluArgAsp 220
QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATAAAAAATCTGCTGTG 962
DB 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022

```

```

DB 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCCGTATCATCCCTTTGATCGCAGTCACCTGAACTGTTCCACATCAATACGTT 1082
DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
DB 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGACATTTATACGATCCTCGTGATAAGCTAACTACTCTCAACAATCTC 1202
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATGCTTTGCTATTATGACATACCTTAACCTGGAAGAGTAGAGGATATCAGCATGAC 1262
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCATACACCGTTTATATGCGCAACGACCGCAAGAGAGAGATGCTAGCTAT 1322
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGCC 1331
DB 381 HisLeuAla 383
RESULT 4
AAI24794
ID AAY24794 standard; protein; 414 AA.
XX
AC AAY24794;
XX
DT 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
DE Streptococcus equisimilis native streptokinase.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nsk; rsk; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN W09931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US026694.
XX
PR 15-DEC-1997; 97US-0069497P.
XX
PA (HARD ) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1999-395183/33.
DR N-PSDB; AAX80492.
XX
PT N-terminally deleted streptokinase.
XX
PS Claim 30; Page 60-61; 73pp; English.
XX
CC The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a

```

CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents native streptokinase (nsk). (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 414 AA;

Alignment Scores:
 Pred. No.: 3,93e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAY24794 (1-414)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGCCCAATTGGTGT 242
 Db 1 IleAlaGlyProGluThrPheLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGACCTGCTCATGAGGAGGAGAGACAGACGAGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCAATTTGCTACTGATGTGCGGATGTCACATAAATCTGAGAAGCTGACTTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
 QY 423 AAGGCTATTCAAGAACAAATGTGCTGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 482
 Db 81 LysAlaIleGlnGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTGCAAGCATGCAACCATTAATGATCGAAACGCGAAGTCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAGATGTTTCGTACCTTGGCGCCACCTGCTCCAGAAATTTTGTGTAAGCGGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACTAAGCTATTGAACACTAGCTATCGGTGACACCATTCACATCTCAAGAAATCTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluAsp 220
 QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

QY 903 TTTACTTACCCTGTTAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTGTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTGAACTGTTCCACCATCAAAATCCTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACGAATTGCTAAAAAGTGAAGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAGGTAAGCTACTCTACAACAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCAGCATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATAACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGGCC 1331
 Db 381 HisLeuAla 383
 RESULT 5
 AAW94664
 ID AAW94664 standard; protein; 414 AA.
 AC AAW94664;
 XX
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis native streptokinase.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 FT
 XX
 FN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR N-PSDB; AAX16632.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 XX

PS Claim 1; Col 7-10; 17pp; English.

XX The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence represents native SK. SK is a secretory protein of
 CC haemolytic streptococcus able to activate human plasminogen (Hplg) to
 CC plasmin (Hplm), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:
 Pred. No.: 3,93e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAW94664 (1-414)

QY 183 ATTGCTGACCTGAGTGGTGTGCTAGACCTGCTCAACACAGCAAGCTTAAAGTGGTGT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
 QY 243 AGCGTGTCTGCTGCTGAGGGGAGCAATCAGACATAGCTTAAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGCACTCTCATGAGGAGAAACAGACAGCAAGCTTAAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACATTTGCTGATAGTGGCGGATGTCACTAACTGAGAAAGCTGACTACTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGCTATTTCAGAACAAATGATCGCTTAAGTCCAGTAAACGACGACTACTTTGAGTTC 482
 Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTTGAAGGATCAACCACTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGTTGCGTAACTTGGCGACCCCACTGCTCCAAAGAAATTTTGTAAAGCGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGCTTAGCCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTACAGACAGTCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATCTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCACAAGCAATTTAAACAAACCCAGGCTATACGATTTATGAACTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATGCTCAGTCAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902

Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGCTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 280
 QY 1023 GAAAGCCGATGATCCCTTTGATCGCAGTCACCTTGAAGCTTCAACATCAAAATGCT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATCAACAGAAATGCTTAAAGTGAGCAGCTTAAACAGCTAGCGAACGTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTTACTCTACAACTCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGCTAGAGGATAATCAGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAAGCTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTAGCC 1331
 Db 381 HisLeuAla 383
 RESULT 6
 ID AAY01556 standard; peptide; 414 AA.
 XX AAY01556;
 AC AAY01556;
 XX 17-OCT-2003 (revised)
 DT 18-JUN-1999 (first entry)
 DE Native streptokinase protein sequence.
 XX Antigenic peptide; streptokinase; streptokinase-specific antibody;
 KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
 KW rheumatic fever.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX WO9908698-A1.
 XX 25-FEB-1999.
 XX 18-AUG-1998; 98WO-US017114.
 XX 18-AUG-1997; 97US-0055911P.
 XX (HARD) HARVARD COLLEGE.
 XX (GEHO) GEN HOSPITAL CORP.
 PI Reed GL, Parhami-Seren B;
 XX WPI; 1999-190113/16.
 XX New polypeptides which bind streptokinase-specific antibodies - useful in
 PT thrombolytic therapy.
 XX Disclosure; Page 12; 44pp; English.
 CC The present sequence represents a native streptokinase. The specification
 CC describes a polypeptide which binds to a streptokinase-specific antibody

CC and prevents the antibody binding to native streptokinase. The
 CC specification also describes a synthetic polypeptide (P1) comprising an
 CC epitope which binds to a streptokinase-specific antibody and reduces
 CC thrombolytic activity of streptokinase. P1 is used in thrombolytic
 CC therapy, and to prevent or treat glomerulonephritis and rheumatic fever.
 CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:

Pred. No.: 3-93e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAY01556 (1-414)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 242
 Db 1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCTTGTCTGCTACTGTTGGGGGAGCAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAAACATCAGACCTGCTCATGTGAGGAGAAACAGACAGCAAGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAACTTCGAGAAGCTGACTTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGCTATTCAAGAACATGATCGCTAAGCTCCACAGTCAACGACGACTACTTTGAGGTC 482
 Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTTTCGACGATCGACCATTTACTGATCGAAACGGCAGGCTCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGGTTCGTAACCTTCGCGACCCCAACCTGTCACAAATTTTTCGTAAGCGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTAGTACGATTCAGACCGTCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATCTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCACAAGCATTTTAAACAAAAACCCAGGCTATACGATTTTATGACCGTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCATCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAATCGGACCAAGCTTATAGATCAATATAAATCTGCTGTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022

Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 280
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAACGAATTCTTAAAGAGTGAGCAGCTCTTTAACAGCTAGCGAACGTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTACACTTCAGAGATTTATACGATCCTCGTGTATAGGCTAAACTCTTACACCAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGTATTATGAGTACTTAACTTAACTGAGAAAGTAGAGGATAATCAGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGAAATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 Db 381 HisLeuAla 383

RESULT 7
 AAY90282
 ID AAY90282 standard; protein; 414 AA.
 XX AC AAY90282;
 XX XX 12-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX DE
 XX S. equisimilis streptokinase.
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin.
 XX OS Streptococcus dysgalactiae subsp. equisimilis.
 XX PN EPI024192-A2.
 XX PD 02-AUG-2000.
 XX PF 23-DEC-1999; 99EP-00310541.
 XX PR 24-DEC-1998; 98IN-DE003825.
 XX PA (COUL) CSIR COUNCIL SCI IND RES.
 XX PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX WPI; 2000-516032/47.
 DR N-PSDB; AAA37633.
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX Example 3; Fig 3; 58pp; English.
 XX This sequence represents the human Streptococcus equisimilis
 CC streptokinase protein sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a

CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 414 AA;

Alignment Scores:
 Pred. No.: 3,93e-178 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 3 Gaps: 0

US-09-940-235-10 (1-1661) x AAY90282 (1-414)

QY 183 ATTGCTGACCTGAGTGGTGTGATGACCGTCTCATCTGTCACACAGCCAAATGGTGT 242
 Db 1 lLeAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
 Db 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 303 GATCTAACATCAGCACTGCTCATGCGAGGAAAGACAGACAGCGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTCAAGCAATTGATCGCTACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
 Db 81 LysAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
 QY 483 ATTGATTTTGCAGCGATGCAACCACTTACTGATCGAAACGCGCAAGGTCTACTTGTGCTGAC 542
 Db 101 lLeAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAlaAsp 120
 QY 543 AAAGATGTTGCGTAACCTTGCAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnGlnGlnGlnGlnGlnGln 140
 QY 603 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAACCAACCAACCAACCAACCAACCAAC 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTACAGCAGGTCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
 Db 181 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 783 GCTCAAGCAACAGCATTTTAAACAAACCAACCAACCAACCAACCAACCAACCAACCAAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCGCTACGATTTTACCAATGGATCAAGAG 902

Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTTACTTACCGTCTTAAATAATCGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrThrValLeuLysLysGly 280
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGAGTCACCTGAAACTGTTCCACATCAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATCAACCAAGTAATGCTAAAAGTGACGACGCTCTTAACAGCTAGCGACGTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTTTATACGATCTCTGCTGATAAGGCTAAACTTACTCTACAACTCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTGTTGTTATGACTATACCTTAACCTGGAAGAGTAGAGGATATCAGCATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAAGCGTATCATACCGTCTTATATGCGCAAGCGACCGAGAGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTAGGCC 1331
 Db 381 HisLeuAla 383
 RESULT 8
 ID AAY24797
 XX AAY24797 standard; protein; 795 AA.
 AC AAY24797;
 XX
 DT 26-AUG-1999 (first entry)
 XX
 DE Streptokinase and maltose binding protein fusion protein.
 XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW rsk; rsk; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 PN WO9931247-A1.
 PD 24-JUN-1999.
 XX
 PF 15-DEC-1998; 98WO-US026694.
 XX
 PR 15-DEC-1997; 97US-0069497P.
 XX
 XX (HARD) HARVARD COLLEGE.
 XX
 PA Reed GL;
 PI
 XX
 DR WPI; 1999-395183/33.
 DR N-PSDB; AAX80497.
 XX
 XX N-terminally deleted streptokinase.
 PT
 XX
 PS Example; Page 48-51; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical

CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents a streptokinase and maltose binding protein fusion protein
 CC from an example of the present invention
 XX
 SQ Sequence 795 AA;

Alignment Scores:

Pred. No.: 4,98e-178 Length: 795
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAY24797 (1-795)

QY 183 ATGCTGGACCTGAGTGGCTGCTGAGACCGTCCATCTCTCAACACAGCCCAATGGTGT 242
 DB 382 IIEAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 401
 QY 243 ACGTTGCTGGTACTGTTGAGGGGGAATCAACACATAGTCTTAATTTTGAATC 302
 DB 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspSerLeuLysPheGluLe 421
 QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGCAAGCTTAAAGTCCAAATCA 362
 DB 422 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 441
 QY 363 AAACATTGCTACTGATGTCGGCGATGTCACATAAAGCTTGAGAAAGCTGACTTACTA 422
 DB 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
 QY 423 AAGGCTATTTCAGAACATGATGCTGCTAAAGTCCACAGTAAACGACGACTACTTGGAGTC 482
 DB 462 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
 QY 483 ATTGATTTTGCAGCGCATGCAACATTTACTGATCGAAACGCAAGGCTACTTTGCTGAC 542
 DB 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
 QY 543 AAAGATGGTTTGGTAACTTTCGCGACCCCAACCTGTCCAAGAATTTTGTGTAAGCGGACAT 602
 DB 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
 QY 603 GTGCGGTTAGACCATATAAGAAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 DB 522 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 541
 QY 663 GAATATCTGTCAGTTTACTCCCTTAAACCCGTGATCGATTCAGCCAGGCTCTCAA 722
 DB 542 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 561
 QY 723 GATCTAAGCTATTGAAACACTAGCTTATCGGTGACACCATCATCTCAAGAAATTACTA 782
 DB 562 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 581

QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCCAGCGCTATACGATTATGAAAGTGAC 842
 DB 582 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgasp 601
 QY 843 TCCTCAATCGTCACCTCATGACAATGACATTTTCCGTAGCATTTTACCAATGATCAAGAG 902
 DB 602 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 621
 QY 903 TTTTACTTACCTGTTTAAAAATCGGCAACAGCTTATAGTCAATAAAAAATCTGCTGTG 962
 DB 622 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 641
 QY 963 AATGAAGAAATAAAACAAACACATGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGG 1022
 DB 642 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 661
 QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATACGTT 1082
 DB 662 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 681
 QY 1083 GATGTCGATACCAAGCAATTCCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAACTAAC 1142
 DB 682 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 701
 QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACTCTC 1202
 DB 702 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 721
 QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAGTAGAGGATAATCACGATGAC 1262
 DB 722 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 741
 QY 1263 ACCAACCGTATCATTAACCGTTTATATGGCGAAGCAGCCGAGGAGAGAGATCTAGCTAT 1322
 DB 742 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 761
 QY 1323 CATTTAGCC 1331
 DB 762 HisLeuAla 764
 RESULT 9
 AAR11829
 ID AAR11829 standard; protein; 531 AA.
 XX
 AC AAR11829;
 DT 08-JUL-1991 (first entry)
 XX
 DE FB-FB-SK fusion conjugate.
 XX
 KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
 KW streptokinase; fusion protein.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 2..61
 FT /label= FB monomer
 FT Peptide 62..119
 FT /label= FB monomer
 FT Peptide 120..531
 FT /label= streptokinase
 XX
 FN US5011686-A.
 XX
 PD 30-APR-1991.
 XX
 PF 15-NOV-1989; 89US-00437769.
 XX
 PR 21-SEP-1987; 87US-00099242.
 XX
 PA (CREA-) CREATIVE BIOMOLEC.
 XX

PI Pang RHL;
 XX WPI; 1991-140198/19.
 DR N-PSDB; AAQ11651.
 XX
 PT Imparting injectable fibrinolytic agent - with affinity for intravascular
 PT thrombus, by linking agent to fibrin binding domain.
 XX
 PS Disclosure; Fig 5; 18pp; English.
 XX
 CC The conjugate comprises an PB-FB dimer linked to streptokinase The PB
 CC fragment has selective affinity for fibrin, low affinity for fibrinogen,
 CC and minimal immunogenicity, imparting thrombus- targeting capability.
 CC See also AAR11821 and AAR11828
 XX
 SQ Sequence 531 AA;
 Alignment Scores:
 Pred. No.: 8.26e-178 Length: 531
 Score: 1981.00 Matches: 382
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 67.40% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-10 (1-1661) x AAR11829 (1-531)
 QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACCAACAGCCCAATGGTGT 242
 DB 118 IleaAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnThrGlnLeuVal 137
 QY 243 AGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
 DB 138 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 157
 QY 303 GATCTAACATCAGCAGCTCTCATGAGGAGAAACAGACAGCAAGCTTAAGTCAAAATCA 362
 DB 158 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGluGlnGlyLeuSerProLysSer 177
 QY 363 AAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
 DB 178 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 197
 QY 423 AAGCTATTCAAGCAATTGATCGCTAACGCTCCAGTAAACGACGACTACTTTGAGTTC 482
 DB 198 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 217
 QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 542
 DB 218 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 237
 QY 543 AAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
 DB 238 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 257
 QY 603 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAACCAACCAACCAACCAACCAACCAAC 662
 DB 258 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 277
 QY 663 GAATATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
 DB 278 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 297
 QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTATCTCAAGAAATACTA 782
 DB 298 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 317
 QY 783 GCTCAAGCAACAGCATTTTAAACCAAAACCAACCAACCAACCAACCAACCAACCAACCAAC 842
 DB 318 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 337
 QY 843 TCCTCAATCGTCACTGACCAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 902

DB 338 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 357
 QY 903 TTTACTTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGCTCG 962
 DB 358 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 377
 QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
 DB 378 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 397
 QY 1023 GAAAGCCGATGATCCCTTTGATCGAGTCACCTTGAACACTGTTTACCATCAATACGTT 1082
 DB 398 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 417
 QY 1083 GATCTCGATACCAACGAATTGCTAAAGAGGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
 DB 418 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 437
 QY 1143 TTAGACTTCAGAGATTTTATACGATCCCTGCTGATGAAGCTAAAGCTTACTCTACCAACATCTC 1202
 DB 438 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 457
 QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
 DB 458 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 477
 QY 1263 ACCAAGCGTATCATACCGTTTATATGGGCAAGCGACCGAGGAGAGATGCTAGCTAT 1322
 DB 478 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 497
 QY 1323 CATTTAGCC 1331
 DB 498 HisLeuAla 500
 RESULT 10
 AAW94665
 ID AAW94665 standard; protein; 414 AA.
 XX
 AC AAW94665;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE Streptococcus equisimilis mutant streptokinase K59E.
 XX
 KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 XX
 PN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Wu H;
 XX
 DR WPI; 1999-189643/16.
 DR N-PSDB; AAX16633.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant

thrombolytic agent.

Claim 4; Col 11-14; 17pp; English.

The present invention describes a mutant streptokinase (SK) polypeptide in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61 segment of the corresponding native SK is replaced by another amino acid. The present sequence is mutant SK K59E. SK is a secretory protein of haemolytic Streptococcus able to activate human plasminogen (HPIg) to plasmin (HPIa), which is a serine protease able to catalyse the hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic agent in the treatment of vascular thromboembolytic symptoms such as acute myocardial infarction. Compared with wild-type SK, the K59E mutant is more resistant to degradation by human plasmin and is more effective both in acting as a fibrinolytic agent and in activating human plasminogen.

Sequence 414 AA;

Alignment Scores:

Pred. No.: 9.38e-178 Length: 414
Score: 1980.00 Matches: 382
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 67.37% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAW94665 (1-414)

183 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 242
1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
243 AGCGTTGCTGTACTGCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
303 GATCTAACATCAGACCTGCTCATGGAGGAAGACAGACGAGCTTAAGTCCAAATCA 362
41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProGluSer 60
363 AAACCATTTGCTACTGATAGTGGCGGATGTCATAAATCTGAGAAAGCTGACTACTA 422
61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
423 AAGCTATTCAAGAAATGATCGCTAACGCTCCACAGTAAACGACACTACTTTGGCTC 482
81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
483 ATTGATTTTGAAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
543 AAAGATGTTGCGTAACTTGGCGACCCCACTGTCCTCAAGAATTTTGTCTAAGCGGACAT 602
121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
603 GTGCGGTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTGTATG 662
141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
663 GAATATCTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACAGGCTCTCAA 722
161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
723 GATACCTAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTTACTA 782
181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
783 GCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGACGTGAC 842
201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluAsp 220
843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCATGGATCAAGAG 902

221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
903 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATATAAATATCTGGTCTG 962
241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
963 AATGAAGAAATAAACAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
1023 GAAAGCGGTATGATCCCTTTGATCGGAGTCCTTGAACCTGTTCAACCATCAATACGTT 1082
281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
1083 GATCTCCATACCAACGAATTGCTAAAAGTCAGAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
1143 TTAGACTTCAGAGATTTATACGATCCGTCGATAGGCTAAACTACTCTACAACTCTC 1202
321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysLeuTyrAsnAsnLeu 340
1203 GATGCTTTTGGTATTATGACTATACCTTAACTCGAAAAGTAGAGGATAATCAGCATGAC 1262
341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1322
361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
1323 CATTTAGGCC 1331
381 HisLeuAla 383
RESULT 11
AAW21728
ID AAW21728 standard; protein; 813 AA.
XX AAW21728;
DT 17-OCT-2003 (revised)
DT 01-OCT-1997 (first entry)
XX Wild type plasminogen-binding fragment of Streptokinase.
XX Plasminogen-binding fragment; streptokinase; degradation; MBP;
XX thrombolytic agent; blood clot; bolus; maltose-binding protein.
OS Streptococcus dysgalactiae subsp. equisimilis.
XX WO9641883-A1.
XX 27-DEC-1996.
XX 07-JUN-1996; 96WO-US0009640.
XX 09-JUN-1995; 95US-00488940.
XX (HARD) HARVARD COLLEGE.
XX Reed GL;
XX WPI; 1997-065469/06.
XX Modified forms of streptokinase resistant to enzymatic cleavage - useful
XX as thrombolytic agents in treating thrombosis and in medical equipment.
XX Example 1; Page 12-13; 65pp; English.
XX This sequence represents the wild type plasminogen-binding fragment of
XX streptokinase. This fragment was used in the design of a modified
XX streptokinase has an in vitro degradation rate at least 2 times slower

CC than that of native streptokinase. Compounds containing modified
 CC streptokinases are specifically used as thrombolytic agents for
 CC dissolving blood clots in vivo in a mammal, preferably at a dose of 20000
 CC U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-
 CC OCT-2003 to standardise OS field)

XX SQ Sequence 813 AA;

Alignment Scores:
 Pred. No.: 2,266-176 Length: 813
 Score: 1966.50 Matches: 382
 Percent Similarity: 99.74% Conservative: 0
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 66.91% Indels: 1
 DB: 2 Gaps: 1

US-09-940-235-10 (1-1661) x AAW21728 (1-813)

Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
 Qy 243 AGCGTTGCTGTGTTGAGGGGACGAATCAAGACATTAGTCTTAATAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 Qy 303 GATCTAAACATCAGCAGCTCTCATGCGAGGAGAAAGACAGCAAGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 Qy 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAAGCTGAGAAAGCTGACTTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 Qy 423 AAGGTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAAACAGCAGTACTTTGAGTC 482
 Db 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 Qy 483 ATTGATTTTGCAGCGATCAACCATTAATGATCGAAACGGCAAGTCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 Qy 543 AAAGATGTTTCGGTAACCTTCGCGACCCCAACCTGTCACAGAATTTTTCGTAACGGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 Qy 603 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAACGAAATCTGTTGATGTG 662
 Db 141 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
 Qy 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTGACAGCGTCTCAAA 722
 Db 160 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
 Qy 723 GATCTAAGCTATTGAACACTAGTATCGTGGTGACACCATCATCATCTCAAGATTAATA 782
 Db 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 199
 Qy 783 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
 Qy 843 TCCTCAATGTCATCATGACAAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
 Db 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
 Qy 903 TTTACTTACCGTAAATCGGACCAAGCTTATAGGATCAATAAAATCTGCTCTG 962
 Db 240 PheThrTyrArgValLysAsnArgGluGlnAlaIleArgIleAsnLysLysSerGlyLeu 259
 Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAGAGGG 1022
 Db 260 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 279

Qy 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATACGTT 1082
 Db 280 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 299
 Qy 1083 GATGTCGATACCAACGAATTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db 300 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 319
 Qy 1143 TTAGACTTCAGAGATTTATACGATCCCTCGTCGATAGGCTAAACTACTCTACACAAATCTC 1202
 Db 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 339
 Qy 1203 GATGCTTTTGGTATTATGGACTATACCTTAACCTCGAAAGTAGAGGATAATCAGCATCAC 1262
 Db 340 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 359
 Qy 1263 ACCAACGCTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAATGCTAGCTAT 1322
 Db 360 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 379
 Qy 1323 CATTTAGCC 1331
 Db 380 HisLeuAla 382

RESULT 12

AAW21726

ID AAW21726 standard; protein; 1194 AA.

XX AC AAW21726;

XX XX 17-OCT-2003 (revised)

DT 01-OCT-1997 (first entry)

XX Streptokinase/maltose binding protein fusion protein, rSK.
 XX Plasminogen-binding fragment; streptokinase; degradation; MBP;
 KW thrombolytic agent; blood clot; bolus; maltose-binding protein.
 XX Streptococcus dysgalactiae subsp. equisimilis.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers

FT Protein 1..381

FT /label= Maltose binding protein

FT /note= "acts as blocking group"

FT 382..1194

FT /label= Streptokinase

XX WO9641883-A1.

XX 27-DEC-1996.

XX 07-JUN-1996; 96WO-US009640.

XX 09-JUN-1995; 95US-00488940.

XX (HARD) HARVARD COLLEGE.

XX Reed GL;

XX WPI; 1997-065469/06.

XX Modified forms of streptokinase resistant to enzymatic cleavage - useful
 as thrombolytic agents in treating thrombosis and in medical equipment.

PS Example 1; Page 11-12; 65pp; English.

XX This sequence represents a fusion protein between maltose-binding protein
 CC and the plasminogen-binding fragment of streptokinase. This fusion
 CC protein was used in the design of modified streptokinase has an in vitro
 CC degradation rate at least 2 times slower than that of native
 CC streptokinase. Compounds containing modified streptokinases are
 CC specifically used as thrombolytic agents for dissolving blood clots in

CC vivo in a mammal, preferably at a dose of 20000 U/kg, optionally as a
CC bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
CC standardise OS field)

SQ Sequence 1194 AA;

Alignment Scores:

Pred. No.: 2,6e-176 Length: 1194
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 66.91% Indels: 1
DB: 2 Gaps: 1

US-09-940-235-10 (1-1661) x AAW21726 (1-1194)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 242
Db 382 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 401
QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 421
QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACAGCAAGGCTTAACTCCAAATCA 362
Db 422 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 441
QY 363 AAAACATTGCTACTGATAGTGGCGGATGTCATAAACTTGAGAAAGCTGACTACTA 422
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
QY 423 AAGGCTATTCAAGACAATTGATCGCTTAAGTCCAGTAAACGACGACTACTTTGGTTC 482
Db 462 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
QY 483 ATTGATTTTGAAGCGCATGCAACCATTTACTGTAGTCAAAACGGCAGGCTCTACTTTGCTGAC 542
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
QY 543 AAGATGTTTGGTAACTTGGTCCGACCCCAACCTGTCCAAAGAATTTTGTCTAAACGCGACAT 602
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
QY 603 GTGCGGCTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTGATGTG 662
Db 522 ValArgValArg---TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTAGTACGATTTTCAGACCGCTCTCAA 722
Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 560
QY 723 GATCTAAGCTATTGAAAACACTAGTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 580
QY 783 GCTCAAGCAAAAGCATTTTAAACAAAACCCAGCTATACGATTTATGAACTGAC 842
Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
QY 843 TCCTCAATCGTCATCATGACATGACATTTTCCGTACGATTTTACCATGGATCAAGAG 902
Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGCTCTG 962
Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
QY 963 AATGAAGAAATAAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1022
Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660
QY 1023 GAAAGCCGTATGATCCCTTTGATGTCAGTCACTTGAAACTGTTCCACCAATAGGTT 1082

Db 661 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 680
QY 1083 GATGTCGATACCAAGCAATTGCTTAAAGTACGAGCTCTTAAACAGCTAGCGAAGTAAAC 1142
Db 681 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTCAAGGCTAAACTACTCTACAACTATCTC 1202
Db 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
QY 1203 GATCCTTTTGGTATTATGGACTATACCTTAACCTGAAAGTAGAGGATAATCACCATGAC 1262
Db 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
QY 1263 ACCAACGCTATCATCAACCGTTTATATGGGCAAGCAGCCGAGAGAGAGAGATGCTAGCTAT 1322
Db 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
QY 1323 CATTTAGCC 1331
Db 761 HisLeuAla 763
RESULT 13
AEB93425
ID AEB93425 standard; protein; 415 AA.
XX AEB93425;
XX 20-OCT-2005 (first entry)
XX Streptococcus equisimilis streptokinase protein SEQ ID NO:1.
XX protein quantitation; streptokinase.
XX Streptococcus dysgalactiae subsp. equisimilis.
XX US2005176085-A1.
XX 11-AUG-2005.
XX 15-NOV-2004; 2004US-00988943.
XX 19-NOV-2003; 2003CU-00000269.
XX (BETA/) BETANCOURT N L H.
PA (DORT/) DORTA-DUQUE J F D C.
PA (PERE/) PEREZ V A B.
PA (VALD/) VALDES J G.
PA (LOPE/) LOPEZ L J G.
PA (PALO/) PALOMARES G R P.
PA (FEYT/) FEYT R P.
PA (GILF/) GIL F M A.
XX Betancourt NLH, Dorta-Duque JFD, Perez VAB, Valdes JG, Lopez LJG;
PI Palomares GRP, Feyt RP, Gil FMA;
XX WPI; 2005-596348/61.
XX Identifying and quantifying proteins in complex mixtures by selectively
PT isolating peptides not containing histidine nor arginine from each
PT protein, and determining the relative concentration of proteins in
PT different samples.
XX Example 3; SEQ ID NO 1; 91pp; English.
XX The invention relates to a method for identifying and quantifying one or
CC more proteins in complex mixtures by selectively isolating peptides not
CC containing histidine nor arginine (NHR peptides) from each protein, and
CC determining the relative concentration of one or more proteins in
CC different samples from the ratio between the areas of estimated
CC theoretical spectra for each NHR peptide labeled with different isotopes
CC in each sample. The method comprises: (a) enzymatically or chemically

hydrolyzing the sample or samples of proteins; (b) chemically modifying alpha and epsilon amino groups (alpha- and epsilon-NH₂) of every peptide obtained in step (a); (c) isolating the NHNR peptides by cation exchange chromatography from the mixture of peptides obtained in step (b); (d) identifying proteins by mass spectrometry analysis of the NHNR peptides obtained in step (c); (e) differential isotopic labeling of protein samples previously to step (a) or during steps (a) or (b) and immediately mixing at least a portion of the samples; and (f) relative quantifying of more proteins in the mixtures of step (e) from the ratio between the areas of estimated theoretical spectra of the pair of NHNR peptides identified in step (d), as well as from the ratio between the areas of the estimated theoretical spectra of fragments from the NHNR peptides, generated in step (d). Also described is a kit for the identification and quantification of proteins in complex mixtures, which comprises the method mentioned above. The method and kit are useful for identifying or quantifying proteins, such as those with vaccinal, therapeutic or diagnostic uses, in complex mixtures. The present sequence represents a streptokinase protein, which is used in an example from the present invention.

Alignment Scores:					
Pred. No.:	3.79e-176	Length:	415		
Score:	1963.00	Matches:	379		
Percent Similarity:	98.96%	Conservative:	1		
Best Local Similarity:	98.70%	Mismatches:	4		
Query Match:	66.79%	Indels:	0		
DB:	9	Gaps:	0		

US-09-940-235-10 (1-1661) x AEB93425 (1-415)

180	ATGATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTTGGTT	2339
Qy		
1	MetIleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
Db		
240	GTTAGCGTTGCTTGTGCTGAGTGGCGAGCAATCAAGACATTAGTCTTAAATTTTGTGAA	299
Qy		
21	ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGlu	40
Db		
300	ATCGATCTAAACATCACGACCTGCTCATGGAGGAAAGACAGACGAAGGCTTAAAGTCCAAAA	359
Qy		
41	IleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLys	60
Db		
360	TCAAAACCATTTGCTACTGATAGTGGCGGATGTCACATAAATCTTGAGAAAGCTGACTTA	419
Qy		
61	SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu	80
Db		
420	CTAAAGGCTATTCAAGAACAAATTGATCGCTACGTCACAGTCCACAGTAAACGACGACTACTTTGAG	479
Qy		
81	LeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspLysPheGlu	100
Db		
480	GTCATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCT	539
Qy		
101	ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValLysPheAla	120
Db		
540	GACAAAGATGCTTCGGTAACTTGGCGGACCCACCTGCTCCAGAAATTTTGTCTAACCGGA	599
Qy		
121	AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly	140
Db		
600	CATGTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCGAAAACTGTGTGAT	659
Qy		
141	HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp	160
Db		
660	GTGGAATATATCTGTACAGTTACTCCCTTAAACCTGTGATACGAAATTCAGACCAGGCTCTC	719
Qy		
161	ValGluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeu	180
Db		
720	AAAGATACTAAGCTATTTCGAACACTAGCTATCGGTGACCAACCATCATCTCTCAAGATTA	779
Qy		
181	LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu	200
Db		
780	CTAGCTCAAGCACAAAGCAATTTTAAACAAAAAACCCAGCGGTATATACGATTTATCAACGT	839
Qy		

PT Reducing immunogenicity of proteins - by modifying the amino acid
PT sequence of the protein to eliminate potential epitopes for T-cells of a
PT given species.

XX Example 6; Fig 28; 77pp; English.

XX The invention relates to a method for the production of non-immunogenic
CC proteins. The method comprises determining at least part of the amino
CC acid sequence of the protein; (b) identifying in the amino acid sequence
CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
CC species; and (c) modifying the amino acid sequence to eliminate at least
CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
CC reduce the immunogenicity of the protein when exposed to the immune
CC system of the given species. A method of analysing a pre-existing protein
CC to predict the basis for immunogenic responses is also provided. The
CC methods can be used particularly for reducing the immunogenicity of
CC immunoglobulins or therapeutic proteins, e.g. Streptokinase (SK). The
CC products can be used for diagnosis and therapy. The present sequence
CC represents the amino acid sequence of the SK protein. (Updated on 17-OCT-
CC 2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:
Pred. No.: 1.12e-175 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x AAW86143 (1-414)

QY 183 ATGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACACACCCCAATGGTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 243 AGCGTTGCTGCTACTGTTAGGCGCGCAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeuPheGluLeu 40
QY 303 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 362
Db 41 AspLeuThr-SerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGGCTATTCAAGACCAATGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGTC 482
Db 81 LysAlaIleGlnGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTGAAGCGATGCAACATTAATGATGCGAAACGCGAGGCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAGATGTTGCGTAACTTGGCCGACCAACCTGTCACAGAAATTTTGTAAAGCGGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGGCGTTAGACCATATAAGAAACCAATACAAACACGAGCAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACGTGTCAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACGAGTCTCAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 723 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200

QY 783 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTATGAACGTGAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 843 TCCTCAATCGTCACATCATGACAATTTTCCGTACGATTTTACCAATGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATAAATAATCTGGCTG 962
Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACAACACTGACTGATCTCTGAGAAATATTAGTCTCTTAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCGGTATGATCCCTTTTGTGCGAGTCAGTCACTTGAAACTGTTCAACATCAATACGTT 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACGAATTGCTAAAGAGTGCAGCTCTTAAACAGCTAGCGACGTAAC 1142
Db 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGAGATTTTATACGATCCCTGATAGGCTAAAGCTAACTACTACAAATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATCCTTTTGGTATTATGGACTATACCTTAACTCGAAAGTAGAGGATAATCAGCATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383
RESULT 15
AAB01295
ID AAB01295 standard; protein; 414 AA.
XX AAB01295;
XX 12-SEP-2003 (revised)
DT 25-SEP-2000 (first entry)
XX
DE Wild type streptokinase from Streptococcus equisimilis.
XX Immunogenicity; immunogen; T cell epitope; T-lymphocyte; drug; vaccine;
KW carrier; diagnosis; therapy; complement; C3; C5; CVF; plasminogen;
KW streptokinase; fibrin; blood clot; thrombolysis; plasmin;
KW myocardial infarction.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS
XX WO200034317-A2.
FN
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-GB004119.
XX
PR 08-DEC-1998; 98GB-00026925.
PR 02-FEB-1999; 99GB-00002139.
XX
PA (BIOV-) BIOVATION LTD.
XX
PI Carr PJ, Adair FS, Hamilton AA, Carter G;
XX
DR WPI; 2000-423372/36.
XX

PT Rendering a protein non-immunogenic or less immunogenic useful in
 PT medicine and in diagnostics involves determining the amino acid sequence
 of the protein, identifying and modifying potential epitopes.

XX Example 2; Fig 5; 42pp; English.

CC Proteins or their fragments can be rendered non-immunogenic or less
 CC immunogenic by identifying one or more potential T cell epitopes and
 CC modifying the sequence to eliminate at least one of these T cell epitopes
 CC to reduce the immunogenicity of the protein when exposed to the immune
 CC system of another organism. The method can be used for producing a
 CC protein with an enzymatic activity which has a beneficial therapeutic
 CC effect, a protein used to convert inactive drugs to its active form
 CC within a living organism, a vaccine, a protein used as a carrier of other
 CC molecule or a protein which binds to other molecules within or introduced
 CC within the living organism in order to alter the bio distribution of
 CC other molecules, such that the protein does not stimulate immune response
 CC in the living organism. The less immunogenic protein is useful in
 CC medicine, diagnosis and in manufacture of a therapeutic or a diagnostic
 CC agent. Streptokinase is produced by certain strains of beta-haemolytic
 CC streptococci. The protein has no inherent enzymatic activity but has
 CC considerable clinical importance owing to its ability to efficiently bind
 CC human plasminogen, potentiating its activation to plasmin and thereby
 CC promoting the dissolution of fibrin filaments in blood clots.
 CC Streptokinase is an effective thrombolytic agent in the treatment of
 CC coronary thrombosis, improving survival and preserving left ventricular
 CC function following myocardial infarction. The native protein is
 CC immunogenic and the production of neutralising antibodies in humans
 CC generally limits the protein to a single use. The new method could
 CC provide a longer therapeutic use for streptokinase. This is the sequence
 CC of the wild type streptokinase. The altered streptokinase sequence is
 CC given in AAB01296. See GENESSEQ records AAB01289-B01302. (Updated on 12-
 CC SEP-2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:

Pred. No.:	1,12e-175	Length:	414
Score:	1958.00	Matches:	378
Percent Similarity:	98.96%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	4
Query Match:	66.62%	Indels:	0
DB:	3	Gaps:	0

US-09-940-235-10 (1-1661) x AAB01295 (1-414)

QY	183	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT	242
DB	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	20
QY	243	ACGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC	302
DB	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	40
QY	303	GATCTAAACATCAGCACCTCTCATGAGGAGACAGACAGACGAGCTTAAGTCCAAATCA	362
DB	41	AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer	60
QY	363	AAACCATTTGCTACTGATAGTGGCGGATGTACATATAAATTCGAGAAAGCTGCTTACTA	422
DB	61	LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	80
QY	423	AAGGCTATTCAAGAACAAATGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGGTC	482
DB	81	LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal	100
QY	483	ATTGATTTTGCAGCGATCAACCATTTACTGATCGAAGCGCAAGGCTCTACTTTGCTGAC	542
DB	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
QY	543	AAAGATGGTTCGGTAACTTGGCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGGACAT	602
DB	121	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	140

Search completed: January 28, 2006, 02:08:17
 Job time : 132.543 secs

QY	603	GTGCGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG	662
DB	141	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	663	GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTTCTCAA	722
DB	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
QY	723	GATCTAAGCTATTGAAAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA	782
DB	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	200
QY	783	GCTCMAGCACAAGACATTTTAAACAAAAACCCAGGCTATACGATTTTATGACGCGTAC	842
DB	201	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	220
QY	843	TCCTCAATCTCTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG	902
DB	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	903	TTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG	962
DB	241	PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu	260
QY	963	AATCMAGAAATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG	1022
DB	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
QY	1023	GAAGAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCAACATCAATACGTT	1082
DB	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
QY	1083	GATGTCGATACCAACGAATTTGCTTAAAGTAGCAGCTCTTAAACAGCTAGCGAACGTAAC	1142
DB	301	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
QY	1143	TTAGCTTCAGAGATTTATACGATCTCTCGTGATAGGCTAAACTACTCTACACAAATCTC	1202
DB	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	340
QY	1203	GATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC	1262
DB	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1263	ACCAACCGTATCATACCCGTTTATATGGCAAGCGACCCGAGAGAGAGAAATGCTAGCTAT	1322
DB	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
QY	1323	CATTAGGCC	1331
DB	381	HisLeuAla	383

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 23.6951 Seconds
(without alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaaccgccagcttagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/USFTO_spool_p/US09940235/runat_27012006_144219_27602/app_query.fasta_1.7708
-DB=Issued_Patents_AA -OFT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CN_1_1_159 @runat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:

- 1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/RE_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfileesi.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	413	2	US-10-360-101-264
2	1984	67.5	414	2	US-09-211-542A-6
3	1984	67.5	414	6	5240845-1
4	1984	67.5	440	1	US-08-560-098A-52
5	1984	67.5	795	2	US-09-211-542A-2
6	1966.5	66.9	813	1	US-08-488-940-3
7	1966.5	66.9	1194	1	US-08-488-940-1
8	1963	66.8	415	1	US-07-854-596B-26
9	1958	66.6	435	1	US-07-854-596B-19
10	1958	66.6	440	1	US-07-854-596B-15
11	1958	66.6	483	1	US-07-854-596B-43
12	1958	66.6	483	1	US-07-854-596B-47

ALIGNMENTS

RESULT 1

US-10-360-101-264

; Sequence 264, Application US/10360101

; Patent No. 6861236

; GENERAL INFORMATION:

; APPLICANT: Moll, Gert N.

; APPLICANT: Leenhouts, Cornelis J.

; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way

; FILE REFERENCE: 2183-5673

; CURRENT APPLICATION NUMBER: US/10/360,101

; CURRENT FILING DATE: 2003-02-07

; PRIOR APPLICATION NUMBER: EP 02077060.8

; PRIOR FILING DATE: 2002-05-24

; NUMBER OF SEQ ID NOS: 309

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 264

; LENGTH: 413

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: sequence of streptokinase

US-10-360-101-264

Alignment Scores:
Pred. No.: 8.87e-186 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-360-101-264 (1-413)

Qy 183 ATTGCTGGACCTGAGTGGCTGCTAGACGCTCCATCTGTCACACACGCAATTCGTTGTT 242

Db 1 ILEAlaGlyProGluTrpLeuLeuAspArgProSerValAsnSerGlnLeuValVal 20

QY 243 AGCGTTGCTGTAAGTCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
 DB |||||||
 QY 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 DB |||||||
 QY 303 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACAGCGTTAAAGTCCAAATCA 362
 DB |||||||
 QY 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 DB |||||||
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAACTTGAGAAAGCTGACTACTA 422
 DB |||||||
 QY 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 DB |||||||
 QY 423 AAGGCTATTCAAGAACCAATTGCTGCTAACGTCACAGTAACACCACTACTTTGAGGTC 482
 DB |||||||
 QY 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 DB |||||||
 QY 483 ATGTATTTTGCAGGATCAACCAATTAAGTCTGAAACGGCAAGCTACTTTGCTGAC 542
 DB |||||||
 QY 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 DB |||||||
 QY 543 AAAGATGTTGCGTAACCTGCGGACCCAACTGTCCTCAAGAAATTTTGTGTAAGCGGACAT 602
 DB |||||||
 QY 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 DB |||||||
 QY 603 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAACCAAGCGAAATCTGTTGATG 662
 DB |||||||
 QY 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 DB |||||||
 QY 663 GAATATCTGTCAGCTTACTCCCTTAAACCCCTGATGACGATTTACAGACAGTCTCAAA 722
 DB |||||||
 QY 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 DB |||||||
 QY 723 GATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCAATCTCAAGAAATTA 782
 DB |||||||
 QY 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 DB |||||||
 QY 783 GCTGAAGCACAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGACATGAC 842
 DB |||||||
 QY 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 DB |||||||
 QY 843 TCCTCAATCGTCATCAGCAATGACATTTTCGTFACGATTTTACCAATGGATCAAGAG 902
 DB |||||||
 QY 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 DB |||||||
 QY 903 TTTACTTACCGTGTAAATAACCGGAACAAGCTTATPAGGATCAATAAAAAATCTGCTG 962
 DB |||||||
 QY 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 DB |||||||
 QY 963 AATGAAGAAATAACCACTGACCTGATCTCTGAGAAATATACGTCCTTAAAAAGGG 1022
 DB |||||||
 QY 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 DB |||||||
 QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCTGCTGAACTGTTACCATCAATCAATCGTT 1082
 DB |||||||
 QY 281 GluLysProTyrAspProPheAspArgSerHisLysLeuLysLeuPheThrIleLysTyrVal 300
 DB |||||||
 QY 1083 GATGTCGATCAACCAAGAAATGCTTAAAGTAGAGCACTCTTAACAGCTAGCGCAACGTAAC 1142
 DB |||||||
 QY 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 DB |||||||
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGTGATAGGCTAAACTACTCTACAAATCTC 1202
 DB |||||||
 QY 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 DB |||||||
 QY 1203 GATGCTTTTGTATTATGACTATACCTTAACCTGGAAGTAGAGGATATACGATGAC 1262
 DB |||||||
 QY 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 DB |||||||
 QY 1263 ACCAACCGGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
 DB |||||||
 QY 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 DB |||||||

QY 1323 CATTAGCC 1331
 DB |||||||
 DB 381 HisLeuAla 383
 DB |||||||
 RESULT 2
 US-09-211-542A-6
 ; Sequence 6, Application US/09211542A
 ; Patent No. 6210667
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Guy L.
 ; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
 ; STREET: 125 Summer Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02110
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/211,542A
 ; FILING DATE: 15-December-1998
 ; CLASSIFICATION: 1653
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/069,497
 ; FILING DATE: 15-December-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Attorney, Strimpel, Harriet M.
 ; REGISTRATION NUMBER: 37,008
 ; REFERENCE/DOCKET NUMBER: 1874/111
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 443-9292
 ; TELEFAX: (617) 443-0004
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 414 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-211-542A-6
 Alignment Scores:
 Pred. No.: 8 88e-186 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-10 (1-1661) x US-09-211-542A-6 (1-414)
 QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATGGTTGTT 242
 DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 303 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACAGCGTTAAAGTCCAAATCA 362
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlyLeuSerProLysSer 60
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAACTTGAGAAAGCTGACTACTA 422
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGGCTATTCAAGAACCAATTGATCGCTAACGTCACAGTAACACCACTACTTTGAGGTC 482

```

Db      81  LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspTyrPheGluVal 100
QY      483  ATTGATTGTTGCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 542
Db      101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY      543  AAGATGTTGCGTAACTTGGCCGACCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT 602
Db      121  LysAspGlySerValThrLeuProThrGlnProValGlnPheLeuLeuSerGlyHis 140
QY      603  GTGCGGTAGACCATATAAGAAACCAACCAATACAAACCAACCAACCAATCTGTTGATGTG 662
Db      141  ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY      663  GAATATACCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGCTCAAA 722
Db      161  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY      723  GATACCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db      181  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY      783  GCTCAAGCAACAGCAATTTTAAACAAACCAACCAACCAACCAACCAATGATGAC 842
Db      201  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY      843  TCCCTCAATCTGCTACTCATGACATGACATTTTCGTACGATTTTACCAATGATCAAGAG 902
Db      221  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY      903  TTTACTACCGTGTAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db      241  PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY      963  AATGAAGAAATAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
Db      261  AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY      1023  GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTACCAATACGTT 1082
Db      281  GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY      1083  GATGTCGATACCAACGAATGCTTAAAGGTAGCAGCTCTTAAACAGCTAGCAGCACTAAC 1142
Db      301  AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY      1143  TTAGACTTCAGAGTTTATACGATCTCTGATAGGCTAAACTTACTCTACAACAATCTC 1202
Db      321  LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY      1203  GATGCTTTTGGTATTATGACATATACCTTAACCTGGAAAGTAGAGGATATACAGATGAC 1262
Db      341  AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY      1263  ACCAACCGTATCATACCGTTTATATGGGCAACGCGACCGCAAGGAGAGATGCTAGCTAT 1322
Db      361  ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY      1323  CATTTAGGCC 1331
Db      381  HisLeuAla 383
```

RESULT 3

5240845-1

; Patent No. 5240845

; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;

; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,

; TSUTOMU

; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS

; NUMBER OF SEQUENCES: 65

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/549,049

; FILING DATE: 06-JUL-1990

; SEQ ID NO: 1:

; LENGTH: 414

5240845-1

Alignment Scores:

Pred. No.:	8.88e-186	Length:	414
Score:	1984.00	Matches:	383
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	67.51%	Indels:	0
Db:	6	Gaps:	0

US-09-940-235-10 (1-1661) x 5240845-1 (1-414)

```

QY      183  ATTCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 242
Db      1  IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY      243  AGCGTTGCTGCTACTGTTGAGGGAGCAATCAAGACATTAAGTCTTAAATTTTTCGAATC 302
Db      21  SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY      303  GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 362
Db      41  AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY      363  AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
Db      61  LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
QY      423  AAGCTATTCAAGAACCAATTCATCGTAACTCCACAGTCCACAGTACGACGACTCTTTGAGT 482
Db      81  LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY      483  ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 542
Db      101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY      543  AAGATGTTGCGTAACTTGGCCGACCAACCTGTCGAAGAAATTTTGTAAAGCGGACAT 602
Db      121  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY      603  GTGCGCTTAGACCATATAAGAAACCAACCAATACAAACCAACCAACCAATCTGTTGATGTG 662
Db      141  ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY      663  GAATATACCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGACAGGCTCAAA 722
Db      161  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY      723  GATACCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db      181  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY      783  GCTCAAGCAACAGCAATTTTAAACAAACCAACCAACCAACCAACCAATGATGAC 842
Db      201  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY      843  TCCTCAATCTGCTACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
Db      221  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY      903  TTTACTTACCGTGTAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db      241  PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY      963  AATGAAGAAATAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
Db      261  AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY      1023  GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTACCAATACGTT 1082
```


Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTCGATACCAAGAAATGCTAAAGCTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluInLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATGCTTTTGGTATTATGACTACTTAACTTAACTGGAAGAGTAGAGGATATCAGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATACCGCTTTATATGGCGAAGCGACCGAAGAGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 Db 381 HisLeuAla 383

RESULT 4

US-08-560-098A-52
 ; Sequence 52, Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WNEWT, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proteins having Fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,098A
 ; FILING DATE: 17-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: P 44 40 892.7
 ; FILING DATE: 17-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42448
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 52:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 440 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-560-098A-52

Alignment Scores:
 Pred. No.: 9,098-186 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0

DB: 1 Gaps: 0
 US-09-940-235-10 (1-1661) x US-08-560-098A-52 (1-440)
 QY 183 ATTGCTGGAGCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATGGTTGTT 242
 Db 27 IleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
 QY 243 AGCGTTCGCTGCTACTGTTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
 QY 303 GATCTTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAAAGCTTTAAAGTCCAAAATCA 362
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACCATTTGCTACTGATGATGCGCGATGTCATAAATCTTGAGAAAGCTGACTTACTA 422
 Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
 QY 423 AAGGCTATTCAAGCAATTTGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 482
 Db 107 LysAlaIleGlnGluInLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
 QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTCTACTTTGCTGAC 542
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 543 AAAGATGGTTCGTAACCTTTGCGACCCCAACCTGTCTCAAGAAATTTTGTAAAGCGGACAT 602
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
 QY 603 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGGAAATCTGTGATGTG 662
 Db 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGTCTCAA 722
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 723 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTTACTA 782
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATAGGATTTATGAACGTGAC 842
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 903 TTTTACTTACCGTGTAAAAATTCGGGAACCAAGCTTTATAGGATCAATAAAAAATCTGTCGTG 962
 Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACCTGAAACCTGTTTCAACATCAATACGTT 1082
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1083 GATGTCGATACCAACGAATTTGCTAAAGGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1142
 Db 327 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTATACGATCTCTGAGTAAGGCTAACTACTCTCAACAATCTC 1202
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATATACGAGTAC 1262

Db 367 AspaAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1263 ACCAACCGTATCAACCGTTTATATGCGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 5

US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998

; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997

; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 443-9292
; TELEFAX: (617) 443-0004

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-211-542A-2

Alignment Scores:
Pred. No.: 1,15e-185 Length: 795
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x US-09-211-542A-2 (1-795)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGCTCCATCTGTCACCAACGCCCAATTTGGTTGT 242
Db 382 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 401
QY 243 ACCGTGCTGCTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 421
QY 303 GATCTAAACATCAGACCTGCTCATGGGAAAGACAGACGAGCTTAAGTCCAAATCA 362

RESULT 6

US-08-488-940-3

; Sequence 3, Application US/08488940

Db 422 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 441
QY 363 AAACCATTTGCTACTGATAGTGGCCGCGATGTACACATAAACTTGAGAAAGCTCAGTACTA 422
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAAACGAGCAGCTACTTTTCAGGTC 482
Db 462 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
QY 483 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGGCAAGGTCTACTTTCTGAC 542
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
QY 543 AAAGATGGTTCGGTAAACCTTCGCGACCCCAACCTGTCGAAGATTTTTCGTAAGCGGACAT 602
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnPheLeuLeuSerGlyHis 521
QY 603 GTGCGCGTTAGACCATATAAAGAAAAACAATACAAAAACCAAGCCAAATCTGTTGATGTG 662
Db 522 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 541
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAAA 722
Db 542 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 561
QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAGAAATTACTA 782
Db 562 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 581
QY 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACAGTGCAC 842
Db 582 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 601
QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 602 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 621
QY 903 TTTACTTACCGTGTATAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 622 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 641
QY 963 AATCAAGAAATAAACAACACCTGATCTCTGAGAAATATTAGTCTCTTAAAAAAGG 1022
Db 642 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLeuGly 661
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGAGTCACCTGCACTTGAACCTGTTCAACCATCAATACGTT 1082
Db 662 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 681
QY 1083 GATCTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTTAACAGCTAGCGAAGCTAAC 1142
Db 682 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 701
QY 1143 TTAGACTTCAGAGATTTTATACGATCCCTGCGATAGCGCTAACTACTCTACACAACTC 1202
Db 702 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 721
QY 1203 GATCTTTTGGTATTATGGACTATACCTTAACTCGAAAAGTAGAGGATAATCAGCATGAC 1262
Db 722 AspaAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 741
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1322
Db 742 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 761
QY 1323 CATTTAGCC 1331
Db 762 HisLeuAla 764

Patent No. 5854049
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,940
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 813 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-940-3

Alignment Scores:
Pred. No.: 6,028-184 Length: 813
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 66.91% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-10 (1-1661) x US-08-488-940-3 (1-813)

QY 183 ATTGCTGACCTGAGTGGCTGTAGACCTGCTCAACCAACAGCAATTTGGTTGTT 242
DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTTGCTGTPACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAGAAAGACAGACAGGCTTAAGTCAAAATCA 362
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTGATGATGCGCGGATGTCACATAAACTCAGAAAGCTGACTTACTA 422
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
QY 423 AAGGCTATTCAAGAACATTTGCTGCTAACGCTCCAGTAACAGCACTACTTTGAGGTC 482
DB 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspThrPheGluVal 100
QY 483 ATTGATTTTGCAGGATGCAACCAATTAAGTATCGAAACGGCAAGTCTACTTTGCTGAC 542
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAlaAsp 120
QY 543 AAAGATGTTGCGTAACCTTGGCGACCCCAACCTGTCTCAAGAAATTTTGTCTAAGCGGACAT 602

DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGGTAGACCATATAAGAAACCAATACCAACCAAGCGAATCTGTTGATCTG 662
DB 141 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
QY 663 GAATATACTGTACAGATTTACTCCCTTAAACCCCTGATCAGCATTTTCAGACAGGCTCAAA 722
DB 160 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 179
QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATTCACATCTCAAGAATTACTA 782
DB 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 199
QY 783 GCTCAAGCACAAAGCATTTTAAACCAAAACCCAGGCTATAGATTATGACGTCAC 842
DB 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
QY 843 TCCTCAATCTCACTCATGCAATGACATTTCCGTACGATTTTACCAATGATCAAGAG 902
DB 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
QY 903 TTTTACTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
DB 240 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysSerGlyLeu 259
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
DB 260 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 279
QY 1023 GAAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTTTCCACATCAAAATAGCTT 1082
DB 280 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 299
QY 1083 GATGTCGATACCAACCAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAAC 1142
DB 300 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 319
QY 1143 TTAGACTTCAGAGATTTATAGATCTCTGCTGATAAGGCTAACTTCTCTCAACAATCTC 1202
DB 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 339
QY 1203 GATGCTTTGTTGTTATGAGTACTTACCTTAACCTGGAAGAGTACAGGATATCAGCATGAC 1262
DB 340 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 359
QY 1263 ACCAACCGTATACACCGTTTATATGGGCAAGCGACCCGAAAGAGAGAATGCTAGCTAT 1322
DB 360 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 379
QY 1323 CATTTAGCC 1331
DB 380 HisLeuAla 382
RESULT 7
US-08-488-940-1
Sequence 1, Application US/08488940
Patent No. 5854049
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,940
 ; FILING DATE: 09-JUN-1995
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 05433/009001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1194 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-488-940-1

Alignment Scores:
 Pred. No.: 7,01e-184 Length: 1194
 Score: 1966.50
 Percent Similarity: 99.74% Matches: 382
 Best Local Similarity: 99.74% Conservative: 0
 Query Match: 66.91% Mismatches: 0
 Indels: 1
 Gaps: 1

US-09-940-235-10 (1-1661) x US-08-488-940-1 (1-1194)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCTGCTCAACACAGCCCAATGGTGT 242
 Db 382 lIeAlaGlyProGluTrpLeuLeuAsePrgProSerValAsnAsnSerGlnLeuVal 401
 QY 243 AGCGTTGTGCTGACTCTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 421
 QY 303 GATCTAACTACGACCTGCTCATGAGGAAAGACAGACGAGCTTAACTCAAAATCA 362
 Db 422 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 441
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAACTGAGAAAGCTGACTACTA 422
 Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
 QY 423 AAGGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 482
 Db 462 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspPyrPheGluVal 481
 QY 483 ATTGATTTTGAAGCGATCAACCACTTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC 542
 Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
 QY 543 AAAGATGGTTCGGTAACTTGGGACCCAACTGTCCAGAAATTTTGTAAACGGGACAT 602
 Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
 QY 603 GTGCGCGGTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db 522 ValArgValArg---TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
 QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGCTCTCAA 722
 Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 560
 QY 723 GATACCTAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 782
 Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 580

QY 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCTATACGATTATTAACAGTGC 842
 Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
 QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAG 902
 Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
 QY 903 TTTACTTACCGTGTAAATAACGGAACCAAGCTTATAGGATCAATAAAAAATCTCGTCTG 962
 Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
 QY 963 AATCAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1022
 Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACATCAATACGTT 1082
 Db 661 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 680
 QY 1083 GATGTCGATACCAACGAATTGCTTAAAGAGTGACGAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
 Db 681 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
 QY 1143 TTAGACTTCAGAGATTTATACGATCCCTGCTGATAGGCTAAACTACTCTACAACAATCTC 1202
 Db 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
 QY 1203 GATGCTTTGCTATTATGACTATACCTTAACTGAGAAAGTAGAGGATAATCAGCATGAC 1262
 Db 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
 QY 1263 ACCAACCGTATCAACCGTTTATATGCGCAAGCGACCGAGGAGAGAGATGCTAGCTAT 1322
 Db 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
 QY 1323 CATTTAGGCC 1331
 Db 761 HisLeuAla 763

RESULT 8

; US-07-854-5968-26
 ; Sequence 26, Application US/078545968
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplowski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,5968
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234

```

; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-26

Alignment Scores:
Pred. No.: 1,02e-183 Length: 415
Score: 1963.00 Matches: 379
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.70% Mismatches: 4
Query Match: 66.73% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x US-07-854-596B-26 (1-415)

QY 180 ATGATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACCAACCAATTTGGTT 239
DB 1 MetIleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 240 GTTAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
DB 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLeuysPhePheGlu 40
QY 300 ATCGATCTAAATCATGACGACCTCTCATGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
DB 41 IleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLys 60
QY 360 TCAAAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
DB 61 SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 80
QY 420 CTAAAGGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAAACGACGACTACTTTGAG 479
DB 81 LeuLysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlu 100
QY 480 GTCAATGATTTTGAAGGATGCAACCATTAATGATGATGATGATGATGATGATGATGATGATG 539
DB 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAla 120
QY 540 GACAAAGATGCTCGGTAACTTCGCGACCTTCGCGACCTTCGCGACCTTCGCGACCTTCGCGAC 599
DB 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly 140
QY 600 CATGTGCGCGTTAGACCATATATAAGAAAAACCAATACAAACCAACGCGAAATCTGTTGAT 659
DB 141 HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 160
QY 660 GTGAATATATCTGTACGTTTACTCCCTTAAACCCCTGATGACGATTCAGACGAGTCTC 719
DB 161 ValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu 180
QY 720 AAAGATATAGCTATTGAAACACTAGTATCGGTGACACCATCATCATCTCAAGATTA 779
DB 181 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY 780 CTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGT 839
DB 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArg 220
QY 840 GACTCCTCAATCGCTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAA 899
DB 221 AspSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240
QY 900 GAGTTTACTACCGTGTAAAAATCGGNAACAAGCTTATAGGATCATATAAAAAATCTGTT 959
DB 241 GluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGly 260
QY 960 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTCGAGAAATATTACGTCCTTAAAAA 1019

```

```

Db 261 LeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLys 280
QY 1020 GGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATAC 1079
Db 281 GlyLulysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 300
QY 1080 GTTGATGTCGATACCAACGAATTTGCTAAAAAGTGCAGCAGCTCTTAAACAGCTAGCGAACGT 1139
Db 301 ValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320
QY 1140 AACTTAGACTTCAGAGATTTATACGATCTCTCGTGAATAAGGCTAACTACTCTACAAAT 1199
Db 321 AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 340
QY 1200 CTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGAT 1259
Db 341 LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 360
QY 1260 GACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGC 1319
Db 361 AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 380
QY 1320 TATCATTTAGCC 1331
Db 381 TyrHisLeuAla 384

RESULT 9
US-07-854-596B-19
; Sequence 19, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-19

Alignment Scores:
Pred. No.: 3.2e-183 Length: 435
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4

```

Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x US-07-854-596B-19 (1-435)
QY 183 ATTGCTGGAGCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAGCCCAATTCGGTTGTT 242
DB 22 ILEAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 41
QY 243 AGCGTTGCTGGTACTGTTGAGGGGCGAATCAACAGACATTAGTCTTAATATTTTGAATC 302
DB 42 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 61
QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 362
DB 62 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 81
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCATAAATTTGAGAAGCTGACTTACTA 422
DB 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 101
QY 423 AAGGCTATTCAAGCAATGTAGTGGCTTAACGTCACAGTAACGACGACTACTTTGAGTC 482
DB 102 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 121
QY 483 ATTGATTTTGAAGCGATGCAACCATTTACTGTGAAACGGCAAGGTCTACTTTGCTGAC 542
DB 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 141
QY 543 AAGATGTTGCTGGTAACCTTGGCCGACCCAACTGTCGAAGAAATTTTGTAAAGCGACAT 602
DB 142 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 161
QY 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
DB 162 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 181
QY 663 GAATATACGTGACAGTTTACTCCCTTAAACCCGTGATGACCATTTTCAGCCAGGCTCAAA 722
DB 182 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
QY 723 -GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
DB 202 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 221
QY 783 GCTCAAGCAAGCAATTTTAAACAAACCCAGCGGTATACGATTTATGAACGTGAC 842
DB 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 241
QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
DB 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 261
QY 903 TTTACTTACCGTTTAAATAATCGGAACAGCTTATAGGATCAATAAAAAATCTGTCGTG 962
DB 262 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 281
QY 963 AATGAAGAAATAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
DB 282 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 301
QY 1023 GAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAAACTGTTCCACCATCAATACGTT 1082
DB 302 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
QY 1083 GATGTCGATACCAAGCAATTTGCTTAAAGTAGACGACGCTCTTACAGCTAGCGACGTAA 1142
DB 322 AspValAsnThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 341
QY 1143 TTAGACTTCAGAGATTATACGATCTCGGTGATAAGCTAAACTACTCTACAAACAATCTC 1202
DB 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 361

QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACCTGAAAAAGTAGAGATAATCACCATGAC 1262
DB 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 381
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAGGAGAGAATGCTAGCTAT 1322
DB 382 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 401
QY 1323 CATTTAGCC 1331
DB 402 HisLeuAla 404
RESULT 10
US-07-854-596B-15
; Sequence 15: Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 15:
; MOLECULE TYPE: protein
; LENGTH: 440 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; Alignment Scores:
Pred. No.: 3,22e-183 Length: 440
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x US-07-854-596B-15 (1-440)
QY 183 ATTGCTGGAGCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAGCCCAATTCGGTTGTT 242
DB 27 ILEAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAACAGACATTAGTCTTAATATTTTGAATC 302
DB 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 66
QY 303 GATCTAACATCAGCACTGCTCATGGAGGAAAGACAGAGCAAGCGCTTAAGTCCAAATCA 362

Db 67 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTCGAGAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGGTC 482
Db 107 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
QY 483 ATTGATTTTGCAGCGATCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 543 AAAGATGTTTCGGTAACCTTGGGACCCAACTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 603 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTC 662
Db 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
QY 723 GATACAGCTATTGAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 782
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCCAGGCTATACGATTTTATGACGTCGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY 843 TCCTCAATCGTCATCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAATAACCGGAACAAAGCTTATAGGATCAATAAAAAATCTGCTCTG 962
Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 286
QY 963 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTAGCTCTTAAAAAGGG 1022
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
QY 1023 GAAAGCCGCTATGATCCCTTTGATCGCTGCTGAACTGTTTCAACCATCAAAATACGTT 1082
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1083 GATGTCGATACCAACGAATTGCTAAAGTAGCAGCTCTTAAACAGCTAGCGCAACGTAAC 1142
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTATACATCTCTGTATAGGCTAAACTACTCTCAACAATCTC 1202
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1203 GATGCTTTTGGTATTTATGAGCTATACCTTAACTGGAAAGTAGAGGATATATCAGATGAC 1262
Db 367 AspAlaPheGlyIleTetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTAT 1322
Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 11

US-07-854-596B-43

; Sequence 43, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-43
Alignment Scores:
Pred. No.: 3 34e-183 Length: 483
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x US-07-854-596B-43 (1-483)
QY 183 ATGTGCGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 242
Db 70 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 89
QY 243 AGCGTTCTGGTACTCTTGGGGGAGCAATCAGACATTAAGTCTTAATTTTGAATC 302
Db 90 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 109
QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAACAGACAGCAAGCTTAAGTCCAAAATCA 362
Db 110 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 129
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTCGAGAAGCTGACTTACTA 422
Db 130 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 149
QY 423 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGGTC 482
Db 150 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 169
QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
Db 170 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 189

QY 543 AAAGATGTTGCGTAACCTTGCGGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACAT 602
Db 190 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 209
QY 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCAATCTGTTGATGTG 662
Db 210 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 229
QY 663 GAATATACCTGTCAGTTTACTCCCTTAAACCCCTGATGACCATTTTCAGACGAGTCTCAAA 722
Db 230 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 249
QY 723 GATACTAAGCTATTGAAAACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 782
Db 250 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 269
QY 783 GCTCAAGCAAAAGCATTAAACAAACCAACCCAGGCTATACGATTTATGATCAAGCTGAC 842
Db 270 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 289
QY 843 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 290 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 309
QY 903 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCTG 962
Db 310 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 329
QY 963 AATGAAGAAATAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 330 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 349
QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCAATACGTT 1082
Db 350 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 369
QY 1083 GATGTCGATACCAACCAATGCTTAAAGTGAAGCAGCTCTTAAACAGTACGCAACGTAAC 1142
Db 370 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 389
QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTAAACTACTCTACAAATCTC 1202
Db 390 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 409
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
Db 410 AspAlaPheGlyMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 429
QY 1263 ACCAACCGTATCATACCGTTTATATGGCGCAAGCGACCCGAGGAGAGTATGCTAGCTAT 1322
Db 430 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 449
QY 1323 CATTTAGCC 1331
Db 450 HisLeuAla 452

RESULT 12
US-07-854-596B-47
; Sequence 47, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA

ZIP: 60606
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854.596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-47
Alignment Scores:
Pred. No.: 3,34e-183 Length: 483
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x US-07-854-596B-47 (1-483)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACAACAGCAATGCTGTGTT 242
Db 1 IleAlaGlyProGluTyrPheLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTTGCTGCTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 303 GATCTAATCATCAGACCTGCTCATGAGGAGAAAGACAGACGAGCTTAAGTCCAAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAGCTGACTTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGGCTATTCAAGAACAAATGATCGCTTAACGCTCACAGTAGTAACGACGACTACTTTGAGGTC 482
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTCAAGCGGATGCAACATTAATGATCGAAACCGCAAGGCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAAGATGTTGCGTAAACCTTGGCGACCCCAACCTGTCCAAGAAATTTTGTCAAGCGGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 723 GATACTAAGCTATTGAAAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTACTA 782

```

181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
783 GCTCAGCACAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGACGTCAC 842
201 AlaGlnAlaGlnSerIleLeuLeuLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
843 TCCTCAATCGTCATCATGACATGACATTTCCGTCACGATTTTACCAATGGATCAAGAG 902
221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
903 TTTACTTACCGTGTAAATAACCGGAACAAGCTTATAGGATCAATAAAAAATCTGTCGTG 962
241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluLeuAsnLysLysSerGlyLeu 260
963 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 1022
261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
1023 GAAAGCCGTATGATCCCTTTCATCGCAGTCACTTCAAACTGTTCAACCATCAATACGTT 1082
281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
1083 GATGTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
301 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
1143 TTAGACTTCAGAGATTTATACATCTCTGATAGGCTAAACTACTCTACAAATCTC 1202
321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC 1262
341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyValGluAspAsnHisAspAsp 360
1263 ACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1322
361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
1323 CATTTAGCC 1331
381 HisLeuAla 383

```

RESULT 13

```

US-07-854-596B-28
; Sequence 28, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-28

Alignment Scores:
Pred. No.: 3 38e-183 Length: 499
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x US-07-854-596B-28 (1-499)
QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCCAATGTTGTT 242
Db 86 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 105
QY 243 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 302
Db 106 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 125
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAAAGCTTAAAGTCCAAAATCA 362
Db 126 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 145
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAACTTCGAGAAGCTGACTTACTA 422
Db 146 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu 165
QY 423 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGTGC 482
Db 166 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 185
QY 483 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 542
Db 186 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 205
QY 543 AAAGATGTTTCGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 602
Db 206 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 225
QY 603 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 226 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 245
QY 663 GAATATACCTGACATTTACTCCCTTAAACCTGATGACGATTTTCAGACGAGCTCTCAA 722
Db 246 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 265
QY 723 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 266 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 285
QY 783 GCTCAAGCACAAGCATTTTAAACAAAACCAAGGCTATACGATTTATGACGTCAC 842
Db 286 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 305
QY 843 TCCTCAATCGTCATCATGACATGACATTTCCGTCACGATTTTACCAATGGATCAAGAG 902
Db 306 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 325
QY 903 TTTACTTACCGTGTAAATAACCGGAACAAGCTTATAGGATCAATAAAAAATCTGTCGTG 962
Db 326 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluLeuAsnLysLysSerGlyLeu 345

```

QY 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 346 AsnGluGluLeuAsnAsnThrAspLeuLeuSerGluLysTyrValLeuLysLysGly 365
QY 1023 GAAAGCGGTATGATCCCTTTGATCGGACGACCTTGAACCTGTTCCACCAATAGCTT 1082
Db 366 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 385
QY 1083 GATGTCGATACCAACGAATTGCTAAAGAGGAGCAGCTCTTAACAGCTAGCGAAGTAAC 1142
Db 386 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 405
QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTTACTCTACAAACAATCTC 1202
Db 406 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 425
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAAAAGTAGAGGATAATCAGCATGAC 1262
Db 426 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 445
QY 1263 ACCAACCGTATCATACCGTTTATATGGCGAAGCGCCGAGGAGAGAAATGCTAGCTAT 1322
Db 446 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 465
QY 1323 CATTTAGGCC 1331
Db 466 HisLeuAla 468
RESULT 14
US-07-854-596B-35
; Sequence 35, Application US/07854596B
; Patent No. 5434b73
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-35
Alignment Scores: 4.19e-183 Length: 859
Pred. No.:

Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x US-07-854-596B-35 (1-859)
QY 183 ATTGCTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACACAGCAATTTGTTCTT 242
Db 22 IleAlaGlyProGluThrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 41
QY 243 AGCGTGTCTGCTACTGTTGAGGGACGAATCAAGACATTAGCTCTTAATATTTTGAATC 302
Db 42 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 61
QY 303 GATCTAATCATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 362
Db 62 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 81
QY 363 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATTTGAGAAGCTGACTTACTA 422
Db 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 101
QY 423 AAGGCTATTCAAGAACCAATTCATCGCTAAACGCTCCACAGTAAACGACGACTATTTGAGTC 482
Db 102 LysAlaIleGlnGluLeuLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 121
QY 483 ATTGATTTTGAACGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 542
Db 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 141
QY 543 AAGATGTTTCGGTAACTTCCCGCCACCACTGTCGAAGAATTTTTCGTAAGCGGACAT 602
Db 142 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 161
QY 603 GTGCGCTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCAAAATCTGTTGATGTG 662
Db 162 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 181
QY 663 GAATATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTTCAGACCAAGGCTCAAA 722
Db 182 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 201
QY 723 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 202 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 221
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCCACCGCTATACGATTTATGAACGTGAC 842
Db 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 241
QY 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGGTACGATTTTCCCAATGGATCAAGAG 902
Db 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 261
QY 903 TTACTTACCGTGTAAAAATTCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGCTG 962
Db 262 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysSerGlyLeu 281
QY 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 282 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 301
QY 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTGCACCTTGAACCTGTTCCACCAATAGCTT 1082
Db 302 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 321
QY 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTAGAGCAGCTCTTAAACAGCTAGCGAAGTAAC 1142
Db 322 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 341
QY 1143 TTAGACTTCAGAGATTTATAGGATCCTCGTGATAGGCTAAACTTACTCTACAAACAATCTC 1202

Db 342 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 361
Qy 1203 GATGCTTTGGTATTATGACTATACCTTAACCTGGAAAGTAGAGGATAATCCGATGAC 1262
Db 362 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 381
Qy 1263 ACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGCAAGGAGAGATGCTAGCTAT 1322
Db 382 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 401
Qy 1323 CATTTAGCC 1331
Db 402 HisLeuAla 404

RESULT 15

US-08-488-940-17
; Sequence 17, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Frazer, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-17

Alignment Scores:
Pred. No.: 6,12e-181 Length: 1194
Score: 1936.50 Matches: 377
Percent Similarity: 98.43% Conservative: 0
Best Local Similarity: 98.43% Mismatches: 5
Query Match: 65.89% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-10 (1-1661) x US-08-488-940-17 (1-1194)

Qy 183 ATTGCTGACCTGAGTGGTGTAGACCGCTCATCTGTCAACCAAGCCGATTTGGTTGT 242
Db 382 IleAlaGlyProGluTrpLeuLeuAspAlaProSerValAsnAsnSerGlnLeuValVal 401
Qy 243 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 302
Db 402 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuAlaPhePheGluIle 421

Qy 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCGCTTAAGTCCAAATCA 362
Db 422 AspLeuThrSerAlaProAlaHisGlyGlyAlaThrGluGlnGlyLeuSerProAlaSer 441
Qy 363 AAACCATTTGCTACTGATAGTGGCGCATGTCCACATAAACTTGAGAAAGCTGACTTACTA 422
Db 442 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 461
Qy 423 AAGGCTATTCAAGAACAAATTGATCGTTAACGTCCACAGTAACGACGACTACTTTGAGGTC 482
Db 462 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 481
Qy 483 ATTGATTTTCCAAGCGATGCAACCATTTACTGATCGAAACCGCAGGCTCTACTTTGCTGAC 542
Db 482 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 501
Qy 543 AAAGATGTTTCGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT 602
Db 502 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 521
Qy 603 GTGCGCTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 522 ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 540
Qy 663 GAATATACTGTACAGTTTACTCCCTTAAACCCGTGATGACGATTTCAGACGAGTCTCAA 722
Db 541 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 560
Qy 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 782
Db 561 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 580
Qy 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCGCTATACGATTTATGACGTCAC 842
Db 581 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 600
Qy 843 TCCTCAATCGTCTCATGACAATGACATTTTCGTACGATTTTACAAATGGATCAAGAG 902
Db 601 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 620
Qy 903 TTTTACTTACCGTGTAAATTCGGACACGCTTATAGGATCAATAAAATCTGCTGCTG 962
Db 621 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 640
Qy 963 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1022
Db 641 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 660
Qy 1023 GAAAAGCCGPTATGATCCCTTTGATCGAGTCATCTTGAACCTGTTTCAACATCAATACGTT 1082
Db 661 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 680
Qy 1083 GATGTCGATACCAACGAATTCCTAAAGGTGACGAGCTCTTAACAGCTAGCGACGTAAC 1142
Db 681 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 700
Qy 1143 TTAGACTTCAGAGATTTATACGATCTCGGTGATAAGGCTAAACTTACTCTACAACTCTC 1202
Db 701 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 720
Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAAGTAGAGGATAATCAGCATGAC 1262
Db 721 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 740
Qy 1263 ACCAACCGTATCATACCGCTTTATATGGGCAAGCGACCCGCAAGGAGAGATGCTAGCTAT 1322
Db 741 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 760
Qy 1323 CATTTAGCC 1331
Db 761 HisLeuAla 763

Search completed: January 28, 2006, 02:41:17
Job time : 42.6951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:51:16 ; Search time 108.622 Seconds
(without alignments)
12778.538 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaacccgcagcgtaccg.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0940235/runat 27012006 144219 27635/app_query.fasta_1.7708
-DB=Published Applications AA Main -QFMT=fasta -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US0940235 -CGN 1 1 805 -runat 27012006 144219 27635
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	413	US-10-360-101-264	Sequence 264, App
2	1984	67.5	414	US-09-940-235-2	Sequence 2, Appli
3	1963	66.8	415	US-10-988-943-1	Sequence 1, Appli
4	1958	66.6	414	US-10-300-215-252	Sequence 252, App
5	1947	66.2	414	US-10-300-215-253	Sequence 253, App
6	1928	65.6	413	US-09-919-703-12	Sequence 12, Appl
7	1794	61.0	440	US-10-474-792-658	Sequence 658, App
8	600	20.4	259	US-09-940-235-4	Sequence 4, Appli
9	600	20.4	2328	US-10-171-311-64	Sequence 64, Appl
10	600	20.4	2328	US-10-236-031B-70	Sequence 70, Appl
11	600	20.4	2328	US-10-374-979-98	Sequence 98, Appl
12	600	20.4	2328	US-10-182-936A-98	Sequence 98, Appl

13	600	20.4	2328	5	US-10-477-238A-677	Sequence 677, App
14	600	20.4	2328	5	US-10-680-287A-677	Sequence 677, App
15	600	20.4	2328	5	US-10-477-173-677	Sequence 677, App
16	600	20.4	2474	5	US-10-450-763-52638	Sequence 52638, A
17	596	20.3	231	3	US-09-925-302-548	Sequence 548, App
18	596	20.3	231	3	US-09-925-302-548	Sequence 548, App
19	596	20.3	463	4	US-10-144-194A-52	Sequence 52, Appl
20	596	20.3	463	5	US-10-491-566-52	Sequence 52, Appl
21	596	20.3	642	4	US-10-741-601-354	Sequence 354, App
22	596	20.3	642	5	US-10-741-600-1066	Sequence 1066, Ap
23	596	20.3	657	4	US-10-741-601-359	Sequence 359, App
24	596	20.3	657	5	US-10-741-600-1072	Sequence 1072, Ap
25	596	20.3	984	4	US-10-741-601-356	Sequence 356, App
26	596	20.3	984	5	US-10-741-600-1069	Sequence 1069, Ap
27	596	20.3	1173	5	US-10-450-763-52634	Sequence 52634, A
28	596	20.3	2220	4	US-10-236-392-4	Sequence 4, Appli
29	596	20.3	2296	4	US-10-741-601-363	Sequence 363, App
30	596	20.3	2296	5	US-10-741-600-1075	Sequence 1075, Ap
31	596	20.3	2355	4	US-10-144-194A-104	Sequence 104, App
32	596	20.3	2355	4	US-10-360-101-235	Sequence 235, App
33	596	20.3	2355	4	US-10-447-161-3	Sequence 3, Appli
34	596	20.3	2355	4	US-10-734-564-94	Sequence 94, Appl
35	596	20.3	2355	4	US-10-741-601-357	Sequence 357, App
36	596	20.3	2355	4	US-10-741-601-366	Sequence 366, App
37	596	20.3	2355	5	US-10-491-566-104	Sequence 104, App
38	596	20.3	2355	5	US-10-741-600-1067	Sequence 1067, Ap
39	596	20.3	2355	5	US-10-741-600-1078	Sequence 1078, App
40	596	20.3	2355	5	US-10-852-335A-147	Sequence 147, App
41	596	20.3	2355	5	US-10-287-436A-436	Sequence 436, App
42	596	20.3	2355	5	US-10-287-436A-1137	Sequence 1137, Ap
43	596	20.3	2355	6	US-11-040-130-28	Sequence 28, Appl
44	596	20.3	2386	3	US-09-961-403-1	Sequence 1, Appli
45	596	20.3	2386	4	US-10-741-601-360	Sequence 360, App

ALIGNMENTS

RESULT 1
US-10-360-101-264
; Sequence 264, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N
; APPLICANT: Leenhouts, Cornelis J
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:
Pred. No.: 1.06e-179 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-360-101-264 (1-413)

Qy 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACGCCAATGGTGTGTT 242
Db 1 lIeAlaGlyProGluIrrpLeuLeuAePargProSerValAenAenSerGInLeuValVal 20

Qy 243 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 Qy 303 GATCTAACATCAGCAGCTCTCATGAGGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 Qy 363 AAACCATTTGCTACTGATAGTCGCGGATGTCACATAAACTTCAGAAAGCTGACTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 Qy 423 AAGGCTATTCAAGAACCAATTTGATCGCTAAACGTCACAGTAACACGACACTACTTTGAGTC 482
 Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 Qy 483 ATTGATTTTGAAGGATCAACCAATCTATGATCGAAACCGGACGCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 Qy 543 AAAGATGTTTCCGTAACTTCGCGACCCAACTGTCACAGAAATTTTGTAAACGGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 Qy 603 GTGCGGTTTAGACCATATAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 Qy 663 GAATATCTGTACGTTTACTCCCTTAAACCCCTGATGACGATTCACAGAGCTCTCAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 Qy 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATCTACTA 782
 Db 181 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 Qy 783 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 Qy 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 Qy 903 TTTACTTACCGTGTAAATCGGAAACAAAGCTTATAGGATCAATAAAAATCTGTCGTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 Qy 963 AATGAAGAAATAAACAACACTGACCTGATCTGTGAAATATTTACGTCCTTAAAAAGGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 Qy 1023 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTCAAACTGTTTCAACCATCAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLysLeuLysPheThrIleLysTyrVal 300
 Qy 1083 GATGTCGATACCAACGAATGTCTAAAGTAGACAGCTCTTAAACAGCTACGCAACCTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 Qy 1143 TTAGACTTCAGAGATTATACCATCTCTGATAGGCTAAAGCTACTCTACACATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATATACCATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 Qy 1263 ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380

Qy 1323 CATTTAGCC 1331
 Db 381 HisLeuAla 383

RESULT 2

US-09-940-235-2
 ; Sequence 2, Application US/09940235
 ; Publication No. US2003059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Streptococcus equisimilis
 ; US-09-940-235-2

Alignment Scores:
 Pred. No.: 1,06e-179 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 3 Gaps: 0

US-09-940-235-10 (1-1661) x US-09-940-235-2 (1-414)

Qy 183 ATTGCTGGACCTGAGTGGCTGTAGACCGCTCATCTGTCAACACAGCCAAATGTTGTT 242
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 Qy 243 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 Qy 303 GATCTAACATCAGCAGCTCTCATGAGGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 Qy 363 AAACCATTTGCTACTGATAGTCGCGGATGTCACATAAACTTCAGAAAGCTGACTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 Qy 423 AAGGCTATTCAAGAACCAATTTGATCGCTAAACGTCACAGTAACACGACACTACTTTGAGTC 482
 Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 Qy 483 ATTGATTTTGAAGGATCAACCAATCTATGATCGAAACCGGACGCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 Qy 543 AAAGATGTTTCCGTAACTTCGCGACCCAACTGTCACAGAAATTTTGTAAACGGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140

QY 603 GTGGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATACCTGTCAGCTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCAAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCACAAGCAATTTAAACAAAAACCCAGCGGTATACGATTTATGAACGTGAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGGATCAAG 902
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 963 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1023 GAAGAGCCGTATGATCCCTTTGATCGCAGTCACCTGAACTGTTCCACCAATACGTT 1082
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1083 GATGTGATACCAACCAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGACGTAAC 1142
 Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
 QY 1143 TTAGACTTCAGAGATTATATACGATCTCTGATGATAAGGCTAAACTACTCTACAACAATCTC 1202
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1203 GATCTTTTGGTATTATGACTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1262
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1263 ACCAACCGTATCATACCGTTTATATGCGCAAGCGCCGAGGAGAGATGCTAGCTAT 1322
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlyGluAsnAlaSerTyr 380
 QY 1323 CATTTAGCC 1331
 Db 381 HisLeuAla 383

RESULT 3

US-10-988-943-1
 ; Sequence 1, Application US/10988943
 ; Publication No. US20050176085A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Center for Genetic Engineering and Biotechnology
 ; TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
 ; TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
 ; TITLE OF INVENTION: COMPLEX MIXTURES
 ; FILE REFERENCE: Proteomics CU2003-269
 ; CURRENT APPLICATION NUMBER: US/10/988,943
 ; CURRENT FILING DATE: 2004-11-15
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 415
 ; TYPE: PRT
 ; ORGANISM: Streptococcus equisimilis
 US-10-988-943-1
 Alignment Scores:

Pred. No.: 1.06e-177 Length: 415
 Score: 1963.00 Matches: 379
 Percent Similarity: 98.96% Conservative: 1
 Best Local Similarity: 98.70% Mismatches: 4
 Query Match: 66.79% Indels: 0
 DB: 5 Gaps: 0
 US-09-940-235-10 (1-1661) x US-10-988-943-1 (1-415)
 QY 180 ATGATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACCAAGCGCAATTCGTT 239
 Db 1 MetIleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 240 GTTAGCGTGTGCTGCTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTTAAATTTTTCGA 299
 Db 21 ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGlu 40
 QY 300 ATCGATCTAACATCACGACCTGCTCATGGAGAAAGACAGACGAAAGCTTAAGTCCAAAA 359
 Db 41 IleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLys 60
 QY 360 TCAAAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTA 419
 Db 61 SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu 80
 QY 420 CTAAGAGCTATTCAAGAACAAATTTGATCGTCAACGTCCACAGTAAACGACGACTACTTTTCAG 479
 Db 81 LeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlu 100
 QY 480 GTCAATTGATTTGCAAGCGATGCAACCACTTACTGATCGAAACGCGAAGTCTACTTTTCT 539
 Db 101 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAla 120
 QY 540 GACAAAGATGTTTCGGTAAACCTTCGCGACCCCACTGTCTCAAGAAATTTTTCGTAAGCGGA 599
 Db 121 AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly 140
 QY 600 CATGTGCGCTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGAT 659
 Db 141 HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 160
 QY 660 GTGCAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGCTCTC 719
 Db 161 ValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu 180
 QY 720 AAGATATCTAAGCTATTGAAACACTAGTATCGGTGACACCATCATCATCTCAAGAAATTA 779
 Db 181 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
 QY 780 CTAGCTCAAGCACAAGCATTTTAAACAAAACCAACCAGCGCTATACGATTTATGAACGT 839
 Db 201 LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArg 220
 QY 840 GACTCTCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATCGATCAA 899
 Db 221 AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 240
 QY 900 GAGTTTACTTACCGTGTAAAAATCGGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGT 959
 Db 241 GluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGly 260
 QY 960 CTGATGAAGAAATAAACCAACACATGACCTGATCTCTGAGAAATATTAGTCTTTAAAAAAA 1019
 Db 261 LeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLys 280
 QY 1020 GGGAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCAACCATCAATAC 1079
 Db 281 GlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 300
 QY 1080 GTTGATGTCGATACCAACGAATTCCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGT 1139
 Db 301 ValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 320

```
Qy 1140 AACTTAGACTTCAGAGATTATACGATCCTCGTGAAGCTAAAGCTAACTACTCTACAACAT 1199
Db 321 AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 340
Qy 1200 CTCGATGCTTTTGGTATTATGACTATACCTTAACCTGGAAAAGTAGAGGATATACGAT 1259
Db 341 LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisasp 360
Qy 1260 GACACCAACCGTATCATCAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGC 1319
Db 361 AspThrAsnArgGileIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 380
Qy 1320 TATCATTTAGCC 1331
Db 381 TyrHisLeuAla 384

RESULT 4
US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252

Alignment Scores:
Pred. No.: 3,18e-177 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.63% Mismatches: 4
Query Match: 66.62% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-300-215-252 (1-414)

Qy 183 ATTGCTGACCTGAGTGGCTGTAGACCTCGTCAACCAAGCCAAATTTGGTTGTT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
Qy 243 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGGAAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
Qy 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCGTCTTAAGTCCAAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
```

RESULT 5

US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1

```
Qy 363 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
Qy 423 AAGCTATTCAAGAACAAATTGATCGCTTAACGCTCCACAGTAAACGACGACTACTTTGAGTC 482
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
Qy 483 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Qy 543 AAAGATGTTTCGGTAACTCTGCCGACCCAACTGCTCCAAAGAAATTTTCTTAAGCGGCAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
Qy 603 GTGCGCTTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy 663 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGAGTCTCAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Qy 723 GATACTAAGCTATTGAAAACACTAGCTATCGTGCACCACTACATCTCTCAAGAAATTACTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
Qy 783 GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCGCTATACGATTTATGAACGTGAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
Qy 843 TCCTCAATCGTCACTCATGACAATGACATTTTCGTCGATGATTTTACCAATGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 903 TTTTACTTTACCGTGTAAATAATCGGGAACACAGCTTATAGGATCAATAAAATCTGGTCTG 962
Db 241 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 260
Qy 963 AATGAAGAAAATAAACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
Qy 1023 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTCAACCATCAATACGTT 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
Qy 1083 GATGTCGATACCAACGAATTGCTTAAAGTAGCAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 301 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
Qy 1143 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAGGCTAAAGCTACTCTACAACAATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
Qy 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGAGAAAGTAGAGGATAATCAGCATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
Qy 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
Qy 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383
```

GENERAL INFORMATION:
 APPLICANT: CARR, Francis Joseph
 APPLICANT: ADAIR, Fiona Suzanne
 APPLICANT: HAMILTON, Anita Anne
 APPLICANT: CARTER, Graham
 TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
 TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
 FILE REFERENCE: MER-104-CON.1
 CURRENT APPLICATION NUMBER: US/10/300,215
 CURRENT FILING DATE: 2002-11-20
 PRIOR APPLICATION NUMBER: US 09/438,136
 PRIOR FILING DATE: 1999-11-10
 PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
 PRIOR FILING DATE: 1998-05-21
 PRIOR APPLICATION NUMBER: GB 9710480.6
 PRIOR FILING DATE: 1997-05-21
 PRIOR APPLICATION NUMBER: GB 9716197.0
 PRIOR FILING DATE: 1997-07-31
 PRIOR APPLICATION NUMBER: GB 9725270.4
 PRIOR FILING DATE: 1997-11-28
 PRIOR APPLICATION NUMBER: GB 9807751.4
 PRIOR FILING DATE: 1998-04-14
 PRIOR APPLICATION NUMBER: US 60/067,235
 PRIOR FILING DATE: 1997-12-02
 NUMBER OF SEQ ID NOS: 254
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 253
 LENGTH: 414
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Modified strep protein

US-10-300-215-253
 Alignment Scores:
 Pred. No.: 3,55e-176 Length: 414
 Score: 1947.00 Matches: 376
 Percent Similarity: 98.69% Conservative: 2
 Best Local Similarity: 98.17% Mismatches: 5
 Query Match: 66.25% Indels: 0
 DB: 4 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-300-215-253 (1-414)

QY	183	ATTGCTGACCTGAGTGGCTGCTAGACCTGCTCAACACAGCCAAATGGTGT	242
Db	1	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	20
QY	243	AGCGTTGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC	302
Db	21	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle	40
QY	303	GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA	362
Db	41	AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	60
QY	363	AAACATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	422
Db	61	LysProPheAlaThrAspSerGlyAlaMetProHisLysLysLysLysLysLysLys	80
QY	423	AAGGCTATTCAAGAAATTTGATCGCTAACGCTCCAGTAACGACGACTCTTGGAGTC	482
Db	81	LysAlaLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	100
QY	483	ATTGATTTTCAAGCGATCAACCATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC	542
Db	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValThrPheAlaAsp	120
QY	543	AAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	602
Db	121	LysAspGlySerValThrLeuProThrGlnProValGlnGlnGlnGlnGlnGlnGln	140
QY	603	GTGCGGTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG	662

Db	141	ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	663	GAATATATCTGTACAGTTTACTCCCTTAAACCCGTGATGACGATTTTACAGACGAGTCTCAA	722
Db	161	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	180
QY	723	GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA	782
Db	181	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	200
QY	783	GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTTTATGAACGTGAC	842
Db	201	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	220
QY	843	TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	902
Db	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
QY	903	TTTACTTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
Db	241	PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu	260
QY	963	AATCAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG	1022
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
QY	1023	GAAGAGCGTATGATCCCTTTGATCGCAGTCACCTTGAACATGTTCCACATCAATAGCTT	1082
Db	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysPheVal	300
QY	1083	GATGTGATACCAACGAATTTGCTAAAGAGGAGCAGCTCTTAACAGCTAGCAGCACTAAC	1142
Db	301	AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
QY	1143	TTAGACTTCAGAGATTTATACGATCTCTCGTGAAGGCTAAAGCTTACTCTCAACAATCTC	1202
Db	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	340
QY	1203	GATGCTTTTGGTATTATGACATACCTTAACTGGAAAGAGTAGAGGATATACGATGAC	1262
Db	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1263	ACCAACGCTATACATACCGTTTATATGGCAAGGACCCCAAGGAGGAGGATGCTAGCTAT	1322
Db	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	380
QY	1323	CATTTAGGCC 1331	
Db	381	HisLeuAla 383	

RESULT 6
 US-09-919-703-12
 Sequence 12, Application US/0919703
 Patent No. US20020165129A1
 GENERAL INFORMATION:
 APPLICANT: Krystal, Gerald
 APPLICANT: Rabkin, Simon W.
 TITLE OF INVENTION: Peptides and Their Use to Ameliorate
 TITLE OF INVENTION: Cell Death
 FILE REFERENCE: 50216/003004
 CURRENT APPLICATION NUMBER: US/09/919,703
 CURRENT FILING DATE: 2001-07-31
 PRIOR APPLICATION NUMBER: US 09/294,457
 PRIOR FILING DATE: 1999-04-19
 PRIOR APPLICATION NUMBER: US 08/759,599
 PRIOR FILING DATE: 1996-12-05
 PRIOR APPLICATION NUMBER: US 60/008,233
 PRIOR FILING DATE: 1995-12-06
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 12
 LENGTH: 413

Qy	483	ATTGATTTTGCACGATGCAACCAATTACTGATCGAAACGCAAGGTCTACTTTGCTGAC	542
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAenGlyLysValTyrPheAlaAsp	146
Qy	543	AAAGATGGTTCCGTAACCTTCGCCGACCCCAACCTGTCCAAAGAAATTTTCTCTAAGCGGACAT	602
Db	147	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis	166
Qy	603	GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG	662
Db	167	ValArgValArgProTyrLysGluLysProValGlnAenGlnAlaLysSerValAspVal	186
Qy	663	GAAATATCTGTACAGTTTACTCCCTTAAACCTGATGACAGATTTTCAGACCAAGGCTCAAA	722
Db	187	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys	206
Qy	723	GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTACTA	782
Db	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	226
Qy	783	GCTCAAGCACAAAGCATTTTAAACAAAAAACCAACCAGCGCTATACGATTATTGAACGTGAC	842
Db	227	AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp	246
Qy	843	TCCCTCAATCGTCACCTCATGACAAAGACATTTTCCGTACGATTTTACCAATCGATCAAGAG	902
Db	247	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	266
Qy	903	TTTACTTACCGTGTATAAAATTCGGGAAACAAGCTTATAGATCAATAAAAAAATCTGGTCTG	962
Db	267	PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle	286
Qy	963	AATCAAGAAATAACAACACATGACCTGATCTCTGAGAAATATATACGCTCTTAAAAAAGGG	1022
Db	287	LysGlyLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly	306
Qy	1023	GAATAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAACCTGTTCACCATCAAAATCGTT	1082
Db	307	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	326
Qy	1083	GATCTCGATACCAACGAATTCTCTAAAAAGTGACGAGCTCTTAAACAGCTAGCGAAACGTAAAC	1142
Db	327	AspValAenThrAenGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346
Qy	1143	TTAAGCTTCAGAGATTTATACGATCTCGTGATAGAGCTAAACCTACTCTACAACAATCTC	1202
Db	347	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	366
Qy	1203	GATCTTTTGGTATTATGGACTATACCTTAACTCGGAAGGTAGAGGATTAATCAGCATGAC	1262
Db	367	AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluLysAsnHisAspLys	386
Qy	1263	ACCAACCGTATCAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGTCTAT	1322
Db	387	AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr	406
Qy	1323	CATTTAGCC	1331
Db	407	HisLeuAla	409

RESULT 8

RESULT 8
US-09-940-235-4

Sequence 4, Application US/09940235

; Publication No. US20030059921A1

; GENERAL INFORMATION:

; APPLICANT: Kumar, Ra

; APPLICANT: Sahni, Girish
: APPLICANT: Roy, Chait

;
: APPLICANT: Roy, Chait
: APPLICANT: Rajagopal.

APPLICANT: RAJAGOPAL, KAMMARA
APPLICANT: Nihalani, Deepak

APPLICANT: Sundaram, Vasudha

APPLICANT: Yadav, Mahavir

```

; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-235-4

Alignment Scores:
Pred. No.: 6.29e-48 Length: 259
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-10 (1-1661) x US-09-940-235-4 (1-259)

```

```

; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-64

```

Alignment Scores:		
Pred. No.:	1.3e-47	Length:
Score:	600.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	20.4%	Indels:
DB:	4	Gaps:
		2328

US-09-940-235-10 (1-1661) x US-10-171-311-64 (1-2328)

QY	1341	CAGCGCAGCAAAATGGTTTCAGCCCACTCCCGGTGCTGTCACTCAAGACAGCCCGGT	1400
Db	5	GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	24
QY	1401	TGTTATCACAATGGAAAACACTATCAGATAAATCAACAGTGGGAGCGGACTACCTAGGT	1460
Db	25	CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly	44
QY	1461	AATGTGTTGTTGTTCTACTGCTTTATGGAGGAAGCGGAGGTTTAACTCCGAAAGTAACCT	1520
Db	45	AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro	64
QY	1521	GAAGCTCAAGACAGACTCTGTTTGCACAGTACACTGGGAACTTACCAGTGGGTGCACCT	1580
Db	65	GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	84
QY	1581	TATGAGCGTCTTAAAGACTCCATGATCTGGACTGTACTCTGCATCGGGGCTGGCGAGGG	1640
Db	85	TyrGluArgProLysAspSerMetIleTipAspCysThrCysIleGlyAlaGlyArgGly	104
QY	1641	AGATAAGCTGTACCATC	1658
Db	105	ArgIleSerCysThrIle	110

RESULT 10

```

US-10-236-031B-70
; Sequence ID, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B0080170265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 70
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-031B-70
Alignment Scores:

```

Pred. No.:	1.3e-47	Length:	2328
Score:	600.00	Matches:	106
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	20.42%	Indels:	0
DB:	4	Gaps:	0

US-09-940-235-10 (1-1661) x US-10-236-031B-70 (1-2328)

Qy	1341	CAGCGCAGCAAAATGGTTTCAGCCGCCAGTCCCGGTGGCTCTCAGTCAAGCAAGCCCGGT	1400
Db	5	GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	24
Qy	1401	TGTTATGACAATCGAAAAACAATTATCAGATAAAATCAACAGTGGGAGCGGACCTACCTAGGT	1460
Db	25	CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnIleTrpGluArgThrTyrLeuGly	44
Qy	1461	AAATGTGTTGTTGTTACTCTGTTATGGAGGAAGCCGAGGTTTTAACTTCGAAAGTAACACT	1520
Db	45	AsnValIleuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro	64
Qy	1521	GAAGCTGAAGAGACTTGCTTTTGACAAGTACACTGGGAACAACATTACCGAGTGGGTGCACCT	1580
Db	65	GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	84
Qy	1581	TATCAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTCGGCCTCGGGGCTGGCGAGGG	1640
Db	85	TyrGluArgProLysAspSerMetIleTleTyrAspCysThrCysIleGlyAlaGlyArgGly	104
Qy	1641	AGAATAAGCTGTACCATC	1658
Db	105	ArgIleSerCysThrIle	110

RESULT 11

```

US-10-374-979-98
; Sequence 98, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-98

```

Alignment Scores:	
Pred. NO.:	1.3e-47
Score:	600.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	20.42%
DB:	4
Length:	2328
Matches:	106
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-940-235-10 (1-1661) x US-10-374-979-98 (1-2328)

Qy	1341	CAGCGCGAGCAAAATGGTTTCAGCCCCAGTCCCCCGTGTCTCAGTCAAAAGCAGCCCGGT	1400
Db	5	GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerIysProGly	24

```
QY 1401 TGTATGACAAATGGAACACATATCATGATAAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY 1461 AATGTGTGGTTGTACTTGTATTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysTyrGlyThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 64
QY 1521 GAAGCTGAAGAGACTTGTCTTTCACAACTGACAGTACCTGCGGAACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110
RESULT 12
US-10-182-936A-98
; Sequence 98, Application US/10182936A
; Publication No.: US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-98
Alignment Scores:
Pred. No.: 1,3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 4 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-182-936A-98 (1-2328)
QY 1341 CAGGCGAGCAAAATGGTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
QY 1401 TGTATGACAAATGGAACACATATCATGATAAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY 1461 AATGTGTGGTTGTACTTGTATTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysTyrGlyThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 64
QY 1521 GAAGCTGAAGAGACTTGTCTTTCACAACTGACAGTACCTGCGGAACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110
```

```
QY 1521 GAAGCTGAAGAGACTTGTCTTTCACAACTGACAGTACCTGCGGAACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
Db 105 ArgIleSerCysThrIle 110
RESULT 13
US-10-477-238A-677
; Sequence 677, Application US/10477238A
; Publication No.: US20040221328A1
; GENERAL INFORMATION:
; APPLICANT: Babi, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-238A-677
Alignment Scores:
Pred. No.: 1,3e-47 Length: 2328
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 5 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-477-238A-677 (1-2328)
QY 1341 CAGGCGAGCAAAATGGTTCAGCCCGAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCGGT 1400
Db 5 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
QY 1401 TGTATGACAAATGGAACACATATCATGATAAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db 25 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY 1461 AATGTGTGGTTGTACTTGTATTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 1520
Db 45 AsnValLeuValCysTyrGlyThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 64
QY 1521 GAAGCTGAAGAGACTTGTCTTTCACAACTGACAGTACCTGCGGAACTTACCGAGTGGGTGACACT 1580
Db 65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY 1581 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGGAGGG 1640
Db 85 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 104
QY 1641 AGAATAAGCTGTACCATC 1658
```



```
Db      105 ArgileSerCysThrIle 110
|||||
RESULT 14
US-10-680-287A-677
; Sequence 677, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-680-287A-677

Alignment Scores:
Pred. No.:      1.3e-47      Length:      2328
Score:          600.00      Matches:    106
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    20.42%      Indels:     0
DB:             5          Gaps:         0

US-09-940-235-10 (1-1661) x US-10-680-287A-677 (1-2328)
QY      1341 CAGGGCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCTCAGTCAAGCAAGCCCGGT 1400
Db      5 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 24
QY      1401 TGTATTGACAATGGAACACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
Db      25 CystyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 44
QY      1461 AATGTGTTGGTTTGTACTTGTATTGAGGAGGAGCGAGGTTTAACTCGGAAAAGTAAACCT 1520
Db      45 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnGlyPheAsnGlySerLysPro 64
QY      1521 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAAACACTTACCGAGTGGGTGACACT 1580
Db      65 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 84
QY      1581 TATGAGCGTCTAAAGACTCCATGATCTGGGACTGTACCTGCTACCTGCGGGGCTGGCGAGGG 1640
Db      85 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 104
QY      1641 AGAATAAGCTGTACCATC 1658
Db      105 ArgileSerCysThrIle 110

Search completed: January 28, 2006, 02:56:50
Job time : 128.622 secs

Db      105 ArgileSerCysThrIle 110
|||||
RESULT 15
US-10-477-173-677
; Sequence 677, Application US/10477173
; Publication No. US2005007069A1
; GENERAL INFORMATION:
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 02:09:11 ; Search time 4.80939 Seconds
(without alignments)
7479.985 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaacccgccagctagcc.....gaataagctgtaccattctaa 1661

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 151242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_27012006_144220_27676/app_query.fasta_1.7708
-DB=Published Applications AA New -QFMT=faetan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USR=US09940235 @CGN_1_1 @runat_27012006_144220_27676
-NCFU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	414	6 US-10-631-558-2	Sequence 2, Appli
2	1928	65.6	413	7 US-11-032-951-12	Sequence 12, Appl
3	600	20.4	259	6 US-10-631-558-4	Sequence 4, Appli
4	596	20.3	300	7 US-11-006-119-31	Sequence 31, Appl
5	596	20.3	642	6 US-10-995-561-631	Sequence 631, App
6	596	20.3	657	6 US-10-995-561-622	Sequence 622, App
7	596	20.3	984	6 US-10-995-561-629	Sequence 629, App
8	596	20.3	2296	6 US-10-995-561-633	Sequence 633, App
9	596	20.3	2355	6 US-10-995-561-623	Sequence 623, App
10	596	20.3	2355	6 US-10-995-561-627	Sequence 627, App

11	596	20.3	2384	6 US-10-821-234-1545	Sequence 1545, Ap
12	596	20.3	2386	6 US-10-995-561-626	Sequence 626, App
13	214	7.3	1341	6 US-10-995-561-621	Sequence 621, App
14	212	7.2	1259	6 US-10-995-561-625	Sequence 625, App
15	211.5	7.2	1286	6 US-10-995-561-628	Sequence 628, App
16	211	7.2	1348	6 US-10-995-561-624	Sequence 624, App
17	208	7.1	693	6 US-10-995-561-632	Sequence 632, App
18	208	7.1	1315	6 US-10-995-561-630	Sequence 630, App
19	199	6.8	847	6 US-10-995-561-634	Sequence 634, App
20	139	4.7	24	7 US-11-006-119-30	Sequence 30, Appli
21	131.5	4.5	2047	7 US-11-013-759-4	Sequence 4, Appli
22	131.5	4.5	2047	7 US-11-013-759-7	Sequence 7, Appli
23	130.5	4.4	1992	7 US-11-013-759-3	Sequence 3, Appli
24	130.5	4.4	1992	7 US-11-013-759-13	Sequence 13, Appli
25	119	4.0	801	7 US-11-174-150-29	Sequence 29, Appl
26	119	4.0	801	7 US-11-124-368A-292	Sequence 292, App
27	119	4.0	1417	7 US-11-052-554A-8	Sequence 8, Appli
28	117	4.0	2314	7 US-11-013-759-11	Sequence 11, Appl
29	116.5	4.0	824	6 US-10-821-234-1008	Sequence 1008, Ap
30	115	3.9	21	7 US-11-032-951-7	Sequence 7, Appli
31	113	3.8	470	6 US-10-485-517-319	Sequence 319, App
32	111	3.8	407	7 US-11-016-564-7	Sequence 7, Appli
33	110.5	3.8	1765	6 US-10-055-877-140	Sequence 140, App
34	110	3.7	1940	6 US-10-055-877-141	Sequence 141, App
35	109	3.7	2890	7 US-11-115-639-31	Sequence 31, Appl
36	109	3.7	2890	7 US-11-115-639-32	Sequence 32, Appl
37	109	3.7	2890	7 US-11-115-639-33	Sequence 33, Appl
38	108.5	3.7	463	6 US-10-733-626-960	Sequence 960, App
39	108	3.7	1647	7 US-11-052-554A-260	Sequence 260, App
40	107.5	3.7	2515	7 US-11-113-424-53	Sequence 53, Appl
41	107	3.6	2710	7 US-11-051-453-41	Sequence 41, Appl
42	106.5	3.6	1193	7 US-11-115-639-19	Sequence 19, Appl
43	106.5	3.6	1193	7 US-11-115-639-20	Sequence 20, Appl
44	106.5	3.6	1193	7 US-11-115-639-21	Sequence 21, Appl
45	106	3.6	876	7 US-11-077-550-82	Sequence 82, Appl

ALIGNMENTS

RESULT 1
US-10-631-558-2
; Sequence 2, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-631-558-2
Alignment Scores:

```

Pred. No.: 9,7e-141 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-10 (1-1661) x US-10-631-558-2 (1-414)

QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCTGCTCAACACAGCCAAATGGTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTTGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluLe 40
QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 362
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCATTTGCTGCTGCTGCTGAGGGGATGTCATAAACTTGAGAAAGCTGACTTACTA 422
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLe 80
QY 423 AAGGCTATTCAAGAAACAATTGATCGCTAACGCTACAGTCCACAGTAACGACGACTTTGAGTGC 482
Db 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 483 ATTGATTTTCAAGCGATCAACCACTTACTGATCGAAACGCGCAAGGTCTACTTTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAAGATGTTGCTGTTAACTTGGCGGACCCAACTGCTCAAGAAATTTTGTGTAAGCGGACAT 602
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 603 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAACGCGAAATCTGTTGATGTC 662
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 663 GAATATACCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACGAGTCTCAAA 722
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTTACTA 782
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
QY 783 GCTCAAGCACAAAGCATTTTAAACAAACACCCAGGCTATACGATTTATGAACTGAC 842
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 843 TCCTCAATCGTCACATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTTACTTACCGGTGTTAAATAACGGAACAAAGCTTATAGGATCAATAAAAAAATCTGCTG 962
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 963 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTTACGCTCTTAAAAAGGG 1022
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCGGTATGATCCCTTTCATCGCAGTCACTTGAACCTGTTCAACCATCAAACTGTT 1082
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACGAATTGCTTAAAGTGTAGCAGCTCTTAAACAGCTAGCGCAACCTAAC 1142
Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320

```

```

QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTACTCTCAACAATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAGGATATACGATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1263 ACCAACCGTATCATACCGCTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383

RESULT 2
US-11-032-951-12
; Sequence 12, Application US/11032951
; Publication No. US20050282741A1
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; APPLICANT: Rabkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003005
; CURRENT APPLICATION NUMBER: US/11/032,951
; PRIOR FILING DATE: 2005-01-11
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/919,703
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-11-032-951-12

Alignment Scores:
Pred. No.: 1,43e-136 Length: 413
Score: 1928.00 Matches: 377
Percent Similarity: 98.95% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 2
Query Match: 65.60% Indels: 2
DB: 7 Gaps: 2

US-09-940-235-10 (1-1661) x US-11-032-951-12 (1-413)

QY 183 ATTGCTGACCTGAGTGGCTGCTAGACCTGCTCAACACAGCCAAATGGTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 243 AGCGTTGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATC 302
Db 21 SerValAlaGlyThrVal---GlyThrAsnGlnAspIleSerLeuLysPhePheGluLe 39
QY 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 362
Db 40 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 59
QY 363 AAACCATTTGCTGCTGCTGCTGAGGGGATGTCATAAACTTGAGAAAGCTGACTTACTA 422
Db 60 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLe 79
QY 423 AAGGCTATTCAAGAAACAATTGATCGCTAACGCTACAGTCCACAGGACGACTTCTGAGTGC 482

```

```

Db      80  LyeAlaileGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 99
QY      483  ATTGATTTCGAAGCGATGCAACATTACTGATCGAAACGCGAGGCTCTACTTTGCTGAC 542
Db      100  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIysValTyrPheAlaAsp 119
QY      543  AAGATGTTTCGGTAACCTTCGCCGACCCCAACCTGTCTCAAGAATTTTTCGTAAGCGGACAT 602
Db      120  LyeAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 139
QY      603  GTCCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db      140  ValArgValArgProTyrIysGluLeuProIleGlnAsnGlnAlaLysSerValAspVal 159
QY      663  GAATATACCTGTCAGATTCTCTCCCTTAAACCCCTGATGACGATTTTCAGACGAGCTCTCAA 722
Db      160  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
QY      723  GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 782
Db      180  LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 199
QY      783  GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGCTATACGATTTATGACGTCGAC 842
Db      200  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
QY      843  TCCTCAATCGTCACTCATGACATGACATTTTCGTCAGATTTTACCAATGGATCAAGAG 902
Db      220  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
QY      903  TTTACTTACCGTGTAAATAATCGGCAACAGCTTATAGGATCAATAAAATCTGGCTCG 962
Db      240  PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysSerGlyLeu 259
QY      963  AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATAT---TACGTCCTTTAAAAA 1019
Db      260  AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluTyrLysTyrValLeuLysLys 279
QY      1020  GGGGAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTACCAATCAATAC 1079
Db      280  GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 299
QY      1080  GTTGATGTCGATACCAACGAATTCCTAAAGTGACGAGCTCTTAACAGCTAGCGACGT 1139
Db      300  ValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArg 319
QY      1140  AACTTACAGCTCAGAGATTTATACGATCTCTCGTGATAAGGCTAAACTTACTCTCAACAAT 1199
Db      320  AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn 339
QY      1200  CTCGATGCTTTTGGTATTATGGAATACTTAACTGGAAGTAGAGGATAATCAACGAT 1259
Db      340  LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp 359
QY      1260  GACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGAGAGAGATGCTAGC 1319
Db      360  AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer 379
QY      1320  TATCAT 1325
Db      380  TyrHis 381

```

RESULT 3

```

US-10-631-558-4
; Sequence 4, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak

```

```

; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-558-4

Alignment Scores:
Pred. No.:          9,22e-38          Length:          259
Score:             600.00             Matches:         106
Percent Similarity: 100.00%           Conservative:    0
Best Local Similarity: 100.00%         Mismatches:     0
Query Match:       20.42%             Indels:         0
DB:                6                 Gaps:          0

US-09-940-235-10 (1-1661) x US-10-631-558-4 (1-259)

QY      1341  CAGCGCAGCAAAATGGTTTCAGCCCACTCCCGGTGGCTGTCACTCAAAAGCAAGCCCGT 1400
Db      1   GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
QY      1401  TGTATCATCAATCGAAACACATCATCATATAACAGTGGAGCGGACCTACCTAGGT 1460
Db      21  CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
QY      1461  AATGTGTGTGTGTGTACTTGTATGTAGAGAGAGCGGAGTGTAACTCGGAAAGTAAACCT 1520
Db      41  AsnValLeuValCysThrCysTyrGlyLysSerArgGlyPheAsnCysGluSerLysPro 60
QY      1521  GAAGCTGAAGAGACTTGTCTTTTGACACAGTACACTGGGAACACTTACCGAGTGGGTGACACT 1580
Db      61  GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
QY      1581  TATGAGCGTCTTAAAGACTCATGATCTGGGACTGTACTACCTGCGGCTGGCGGCGGAGGG 1640
Db      81  TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
QY      1641  AGAATAAGCTGTACCATC 1658
Db      101  ArgIleSerCysThrIle 106

RESULT 4
US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Brian
; APPLICANT: Ndao, Momar
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: CIPHERGEN Biosystems, Inc.
; TITLE OF INVENTION: Serum Biomarkers for Chagas Disease
; FILE REFERENCE: 016866-012130US
; CURRENT APPLICATION NUMBER: US/11/006,119

```

```
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/527,153
; PRIOR FILING DATE: 2003-12-05
; PRIOR APPLICATION NUMBER: US 60/565,093
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US 60/625,519
; PRIOR FILING DATE: 2004-11-06
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
US-11-006-119-31

Alignment Scores:
Pred. No.: 1.86e-37 Length: 300
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: Gaps: 0

US-09-940-235-10 (1-1661) x US-11-006-119-31 (1-300)

QY 1341 CAGGCGCAGCAATGGTTGACGCCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCGGT 1400
Db 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
QY 1401 TGTATTGACAAATGGAACACTATCAGATAAAATCAACAGTGGGAGCGACCTACCTAGT 1460
Db 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
QY 1461 AATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1520
Db 41 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
QY 1521 GAAGCTGAAGAGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1580
Db 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
QY 1581 TATGAGCGTCCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCTGGGCTGGGCGAGG 1640
Db 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
QY 1641 AGAATAAGCTGTACCATC 1658
Db 101 ArgIleSerCysThrIle 106

RESULT 5
US-10-995-561-631
; Sequence 631, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-631

Alignment Scores:
Pred. No.: 2.03e-37 Length: 642
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-631 (1-642)

QY 1341 CAGGCGCAGCAATGGTTGACGCCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCGGT 1400
Db 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAAATGGAACACTATCAGATAAAATCAACAGTGGGAGCGACCTACCTAGT 1460
Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1520
Db 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91

US-09-940-235-10 (1-1661) x US-10-995-561-631 (1-642)

QY 1341 CAGGCGCAGCAATGGTTGACGCCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCGGT 1400
Db 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAAATGGAACACTATCAGATAAAATCAACAGTGGGAGCGACCTACCTAGT 1460
Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1520
Db 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
```

Qy	1521	GAAGCTGAAGAGACTTGGCTTTGACAAGTACACTGGGACAACCTTACCGAGTGGGTGCACACT	1580
Db	92	GluAlaGluGluThrCysPheAspLysIleThrGlyAsnThrTyrArgValGlyAspThr	111
Qy	1581	TATGAGCGTCTAAAGACTCCATGACTCTGGGACTGTACTCTGTCATCGCGGGCTTGGCGCAGGG	1640
Db	112	TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly	131
Qy	1641	AGATAAGCTGTACCATC	1658
Db	132	ArgIleSerCysThrIle	137

RESULT 7

```

US-10-995-561-629
; Sequence 629, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 629
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-629

```

Alignment Scores:		
Pred. No.:	2,14e-37	984
Score:	596.00	105
Percent Similarity:	99.06%	Conservative: 0
Best Local Similarity:	99.06%	Mismatches: 1
Query Match:	20.28%	Indels: 0
DB:	6	Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-629 (1-984)

1341	Qy	CAGCGCAGCAAAATGGTTTCACGCCCCAGTCCCGGTGGCTGTCACTCAAAGCAAGCCCGGT	1400
32	Db	GLNALaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	51
1401	Qy	TGTTATCACAATGGAAGAACACTATCAGATAAATCAACAGTGGGAGCGGACTCTACCTAGGT	1460
52	Db	CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly	71
1461	Qy	AATGTGTGTGGTTGTACTCTCTTATGGAGGAGCCGAGGTTTAACTCGGAAGCTAAACCT	1520
72	Db	AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro	91
1521	Qy	GAAGCTCAAGAGACTCTGTTTGAACAAGTACACTGGGAACACTTACCGAGTGGGTGACACT	1580
92	Db	GluaLaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	111
1581	Qy	TATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTACTCGCATCGGGGTGGCCAGGG	1640
112	Db	TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly	131
1641	Qy	AGAATAAGCTGTATCCATC	1658
132	Db	ArailSerCysThrIle	137

RESULT 8

RESOL 8
US-10-995-561-633
; Sequence 633, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 633
; LENGTH: 2296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-633

```

Alignment Scores:		
Pred. No.:	2,36e-37	Length:
Score:	596.00	Matches:
Percent Similarity:	99.06%	Conservative:
Best Local Similarity:	99.06%	Mismatches:
Query Match:	20.28%	Indels:
DB:	6	Gaps:

US-09-940-235-10 (1-1661) X US-10-995-561-633 (1-2296)

Qy	1341	CAGCGCAGCAAAATGGTTTCAGCCCACTCCCCTGGCTGTCACTCAAGCAAGCCCGGT	1400
Db	32	GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	51
Qy	1401	TGTTATGACAAATCGAAAAACACTATCAGATAAATCAACAGTGGGAGCGGACTCTACCTAGGT	1460
Db	52	CysTyrAspAsnGlyLysHisIstYrGlnIleAsnGlnGlnItrpGluArgThrTyrLeuGly	71
Qy	1461	AATGTGTGTGTTTGTACTCTGTTATGGAGGAAGCCGAGGTTTTAACTCGGAAGCTAAACCT	1520
Db	72	AsnAlaIeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro	91
Qy	1521	GAAGCTGAAGAGACTTGTCTTTTGACAAGTCACTGGGAACTATACCAGTGGGTGCACCT	1580
Db	92	GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	111
Qy	1581	TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGTGGCGCAGGG	1640
Db	112	TyrGluArgProLysAspSerMetIleItrpAspCysThrCysIleGlyAlaGlyArgGly	131
Qy	1641	AGAATAAGCTGTACCATTC	1658
Db	132	ArgIleSerCysThrIle	137

RESULT 9

```

US-10-995-561-623
; Sequence 623, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 623
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-623

Alignment Scores:
Pred. No.: 2.36e-37 Length: 2355
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1

```

Query Match: 20.28% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-995-561-623 (1-2355)
QY 1341 CAGGCGCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCGGT 1400
DB 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAATGGAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
DB 52 CysTyrAspAsnGlyshisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAAACCT 1520
DB 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
QY 1521 GAAGCTGAAGACACTGCTTTCACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 1580
DB 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCTAAAGACTCCATGATCTGGAGTGTACCTGCATCGGGGTGGCGAGGG 1640
DB 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
DB 132 ArgIleSerCysThrIle 137
RESULT 10
US-10-995-561-627
; Sequence 627, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-627
Alignment Scores:
Pred. No.: 2,368-37 Length: 2355
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-995-561-627 (1-2355)
QY 1341 CAGGCGCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCGGT 1400
DB 32 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATTGACAATGGAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
DB 52 CysTyrAspAsnGlyshisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGTTGTTGTTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAAACCT 1520
DB 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
QY 1521 GAAGCTGAAGACACTGCTTTCACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 1580
DB 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCTAAAGACTCCATGATCTGGAGTGTACCTGCATCGGGGTGGCGAGGG 1640
DB 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
DB 132 ArgIleSerCysThrIle 137

DB 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCTAAAGACTCCATGATCTGGAGTGTACCTGCATCGGGGTGGCGAGGG 1640
DB 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
DB 132 ArgIleSerCysThrIle 137
RESULT 11
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC SEQ_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1545
Alignment Scores:
Pred. No.: 2,378-37 Length: 2384
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-10 (1-1661) x US-10-821-234-1545 (1-2384)
QY 1341 CAGGCGCAGCAATGTTTCAGCCCGAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCGGT 1400
DB 61 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 80
QY 1401 TGTATTGACAATGGAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
DB 81 CysTyrAspAsnGlyshisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 100
QY 1461 AATGTTGTTGTTGTTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAAACCT 1520
DB 101 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 120
QY 1521 GAAGCTGAAGACACTGCTTTCACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 1580
DB 121 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 140
QY 1581 TATGAGCGTCTAAAGACTCCATGATCTGGAGTGTACCTGCATCGGGGTGGCGAGGG 1640
DB 141 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 160
QY 1641 AGAATAAGCTGTACCATC 1658
DB 161 ArgIleSerCysThrIle 166
RESULT 12
US-10-995-561-626
; Sequence 626, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 626
; LENGTH: 2386
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-626

Alignment Scores:
Pred. No.: 2,37e-37 Length: 2386
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 20.28% Indels: 0
DB: Gaps: 0

US-09-940-235-10 (1-1661) x US-10-995-561-626 (1-2386)
QY 1341 CAGGCGCAGCAATGTTGAGCCAGTCCCGTGGCTGCTCAGTCAAGCAAGCCCGT 1400
DB 32 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATGACAATGGAAACATCATCATGATAATCAACAGTGGGAGCGGACTACCTAGT 1460
DB 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnInTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTTGCTGTTGTTACTTGTATGAGGAGCGGAGGTTTAACTCGGAAAGTAAACCT 1520
DB 72 AsnAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 91
QY 1521 GAAGCTGAAGAGACTTGTCTTGAACAGTACACTGGGAACACTTACCGAGTGGGTGACACT 1580
DB 92 GluAlaGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCTCTTAAGACTCCATGATCTGGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1640
DB 112 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658
DB 132 ArgIleSerCysThrIle 137

RESULT 13
US-10-995-561-621
; Sequence 621, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 1341
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-621

Alignment Scores:
Pred. No.: 5.95e-09 Length: 1341
Score: 214.00 Matches: 133
Percent Similarity: 32.70% Conservative: 73
Best Local Similarity: 21.11% Mismatches: 200

Query Match: 7.28% Indels: 224
DB: Gaps: 29
US-09-940-235-10 (1-1661) x US-10-995-561-621 (1-1341)
QY 15 CTAGCCGGTCTCTCAACGACGAGGACGATCATCGCGCCGCGGCGGACGCCAAGCC 74
DB 789 LeuIleGlyThrGlnSerThrAlaIleProAlaProThrAspLeuLysPheThrGlnVal 808
QY 75 TGCCCGAGATCTCGATCCCGGAA-----ATTAAATACGACTCATATAGG 119
DB 809 ThrProThrSerLeuSerAlaGlnTrpThrProProAsnValGlnLeuThrGlyTyrArg 828
QY 120 GAGACCACACGCGTTTCCCTCTAGAAATAATTTTGTAACTTTAAGAGGAGATATACC 179
DB 829 ValArgValThrProLys----- 834
QY 180 ATGATTGCTGACCTCGAGTGGCTGTAGACGCTCATCTGCAACACAGCAATGTT 239
DB 835 GluLysThrGlyProMetLysGluIleAsnLeuAlaProAspSerSer----- 851
QY 240 GTTAGCGTTGCTGCTGCTGAGGGGAGCAATCAAGACATAGTCTTAAATTTTTCGA 299
DB 852 ValValValSerGlyLeuMetValAlaThrLysTyrGluValSerValTyrAlaLeuLys 871
QY 300 ATCGATCTAAACATCAGCAGCTGCTCATGGAGGAAAGACAGAG---CAAGGCTTAAGTCCA 356
DB 872 AspThrLeuThrSerArgProAlaGlnGlyValValThrThrLeuGluAsnValSerPro 891
QY 357 AAATCAAAACCATTTGCTACTGATGATGGCGCGATGTACATAAATTTGAGAAAGCTGAC 416
DB 892 ProArgArgAlaArgValThrAsp----- 899
QY 417 TTACTAAGGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAACGACGACTCTTT 476
DB 899 ----- 899
QY 477 GAGTCTATTGATTTTGAAGCGATGCAACCACTTACTGATCGAAGCGCAAGGTCTACTTT 536
DB 900 -----AlaThrGluThrThrIleThrIle----- 907
QY 537 GCTGACAAAGATGTTGCGTAACCTTGCAGCAACCACTGCTCCAAAGATTTTGTCTAAGC 596
DB 908 -----SerTyrArgThrLysThrGluThrIleThrGlyPheGlnValAsp 922
QY 597 GGACATGTGCGGTTAGACCATATAAGAAAAACAATACAAACCAAGCCAAATCTGTT 656
DB 923 Ala-----ValProAlaAsnGlyGlnThrProIleGlnArgThrIleLysProAsp 939
QY 657 GATGTGGAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCAGGT 716
DB 940 ValArgSerTyrThrIle-----ThrGlyLeuGlnProGlyThrAspTyrLys----- 955
QY 717 CTCAAAGATATAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAA 776
DB 956 -----IleTyrLeuTyrThrLeuAsnAspAsnAlaArgSerSerPro 969
QY 777 TTACTAGCTCAAGCA----- 791
DB 970 ValValIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheLeuAlaThr 989
QY 792 -----CAAAGCATTTTAAACAAAACCAACCCA-----GGCTATACG 827
DB 990 ThrProAsnSerLeuLeuValSerTrpGlnProProArgAlaArgIleThrGlyTyrIle 1009
QY 828 ATT---TATGAAGTGTACTCTCAATCGTCACTCATGACATGACATTTTTCGTCAGCAT 884
DB 1010 IleLysTyrGluLysProGlySerProPro-----ArgGluVal 1022
QY 885 TTACCAATGGATCAA----- 899
DB 1023 ValProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGluProGlyThr 1042


```

QY 792 -----CAAGCATTTTAAACAAAACCCCA-----GGCTATACG 827
Db 990 ThrProAsnSerLeuValSerTrpGlnProProArgAlaArgIleThrGlyTyrIle 1009
QY 828 ATT---TATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATT 884
Db 1010 IleLysTyrGluYsProGlySerProPro-----ArgGluVal 1022
QY 885 TTACCAATGGATCA----- 899
Db 1023 ValProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGluProGlyThr 1042
QY 900 GAGTTTACTTACCGTGT-----AAAAATCGGGAACAAGCTTATAGATCAAT 947
Db 1043 GluTyrThrIleTyrValIleAlaLeuLysAsnGlnLysSerGluProLeuIleGly 1062
QY 948 AAAAATCTGGTCTGAATGAAGAATAAACAAC---ACTGACCTGATCTCTGAGAAATAT 1004
Db 1063 ArgLysLysThrValAsnGluGlyLeuAsnGlnProThrAsp----- 1076
QY 1005 TACGTCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACATG 1064
Db 1077 -----AspSerCysPheAspProTyrThrValSerHis----- 1087
QY 1065 TTCACCATCAATPACGTTGATGTCGATACCAAGAAATGTCTAAAGAGTACGACCTCTTA 1124
Db 1088 -----TyrAlaValGlyAspGluTrpGluArgMet---SerGluSerGlyPhe 1102
QY 1125 ACAGTAGCGAAGTAACTTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTAAA 1184
Db 1103 LysLeuLeuCysGlnCysLeuGlyPheGlySerGlyHisPheArgCysAspSerSerArg 1122
QY 1185 CTACTCTACAAATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTA 1244
Db 1123 TrpCysHisAspAsn-----GlyVal---AsnTyr-----LysIle 1133
QY 1245 GAGGATAATCAGCATGACACCAACCGGTATCATACCGCTTTATATGGCAAGCGACCGAA 1304
Db 1134 GlyGluYsTrpAsp----- 1140
QY 1305 GAGAGAAATGCT-----AGCTATCATTTAGCCGGTGGTGTGTCAGCGCAGCAAAATG 1355
Db 1141 GlyGluAsnGlyGlnMetMetSerCysThrCysLeuGlyAsnGlyLysGlyGluPheLys 1160
QY 1356 GTTCAGCCCGCAGTCCCGGTGGTGTGTCAGTCAAGCAAGCCCGGTGTTATGACAATGGA 1415
Db 1161 CysAspProHisGluAlaThr-----CysTyrAspAspGly 1172
QY 1416 AAACACTATCAGATAAATCAACAGTGGGAGCGACCTTACCTAGTAAATGTGTGTTGT 1475
Db 1173 LysThrTyrHisValGlyGluGlnTrpGlnLysGluTyrLeuGlyAlaIleCysSerCys 1192
QY 1476 ACTTGTATTGAGGAAGCCGAGTTTAACTGCGAAAGT-----AAA 1517
Db 1193 ThrCysPheGlyGlyGlnArgGlyTyrArgCysAspAsnCysArgArgProGlyGlyGlu 1212
QY 1518 CTGAAGCTGAAGAGACTGCTTTTGACAAGTACACTGGGAACACTTAC 1565
Db 1213 ProSerProGluGlyThr-----ThrGlyGlnSerTyr 1223

```

RESULT 15

```

US-10-995-561-628
; Sequence 628, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702

```

```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 628
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-628

Alignment Scores:
Pred. NO.: 9.09e-09 Length: 1286
Score: 211.50 Matches: 126
Percent Similarity: 33.28% Conservative: 72
Best Local Similarity: 21.18% Mismatches: 188
Query Match: 7.20% Indels: 209
DB: 6 Gaps: 27

```

```

US-09-940-235-10 (1-1661) x US-10-995-561-628 (1-1286)

QY 15 CTACCCGGTCTCTCAACGACGAGGAGCAGCATCGCGACCGGTGGCCGAGCAACGCG 74
Db 789 LeuIleGlyThrGlnSerThrAlaIleProAlaProThrAspLeuLysPheThrGlnVal 808
QY 75 TGCCCGAGATCTCGATCCCGGAA-----ATTAATACGACTCCTATAGG 119
Db 809 ThrProThrSerLeuSerAlaGlnTrpThrProProAsnValGlnLeuThrGlyTyrArg 828
QY 120 GAGACCAACACGGTTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGGAGATATACC 179
Db 829 ValArgValThrProLys----- 834
QY 180 ATGATGCTGCACTGAGTGGCTGTAGACGCTCATCTGTCAACACAGCAATGTTGTT 239
Db 835 GluLysThrGlyProMetLysGluIleAsnLeuAlaProAspSerSerSer- 851
QY 240 GTTAGCGTGTCTGCTACTGTGTGAGGGGACGAAATCAAGACATTAAGTCTTAAATTTTGA 299
Db 852 ValValValSerGlyLeuMetValAlaThrLysTyrGluValSerValTyrAlaLeuLys 871
QY 300 ATCGATCTTAACATCAGCAGCTGCTCATGGAGGAAAGCAGAG---CAAGGCTTAAGTCCA 356
Db 872 AspThrLeuThrSerArgProAlaGlnGlyValValThrLeuGluAsnValSerPro 891
QY 357 AAATCAAAACCATTTGCTACTGTAGTGGCGCGATGTCCACATAAACTTGAGAAAGCTGAC 416
Db 892 ProArgAlaArgValThrAsp----- 899
QY 417 TTACTAAAGGCTATTCAAGAACAAATGATCGCTAAGTCCACAGTAAAGCAGCAGTACTTT 476
Db 899 ----- 899
QY 477 GAGTCTATTGATTTTCAAGCGATGCAACCACTTACTGATCGAAACGGCAAGGTCTACTTT 536
Db 900 -----AlaThrGluThrIleThrIle----- 907
QY 537 GCTGACAAAGATGTTCCGTAACCTTCCGCAACCAACCTGCTCAAGAAATTTTGTCAAGC 596
Db 908 -----SerTyrArgThrLysThrLysThrLysThrLysThrLysThrLysThrLys 922
QY 597 GGACATGTGCGGTAGACCATATATAAGAAAAACAATACAAAAACCAAGCGAAATCTGTT 656
Db 923 Ala-----ValProAlaAsnGlyGlnThrProIleGlnArgThrIleLysProAsp 939
QY 657 GATGTGGAATATACTGTACAGTTTACTCCCTTAAACCTGTATGAGTTCAGACACCGGT 716
Db 940 ValArgSerTyrThrIle-----ThrGlyLeuGlnProGlyThrAspTyrLys----- 955
QY 717 CTCAAGATACTAAGCTATTGAAACACACTAGCTATCGTGTGACACCATCACATCTCAAGAA 776
Db 956 -----IleTyrLeuTyrThrLeuAsnAspAsnAlaArgSerSerPro 969
QY 777 TTAAGTCTCAAGCA----- 791
Db 970 ValValIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheLeuAlaThr 989

```

```
QY 792 -----CAAAGCATTTTAAACAAAAACCCCA-----GGCTATACG 827
Db |||:|||||
QY 990 ThrProAsnSerLeuLeuValSerTrpGlnProArgAlaArgIleThrGlyTyrIle 1009
Db |||:|||||
QY 828 ATT-----TATGAAGTCATCTCCTCAATCGTCACATGACACATGATTTCCGTACGATT 884
Db |||:|||||
QY 1010 IleLysTyrGluLysProGlySerProPro-----ArgGluVal 1022
Db |||:|||||
QY 885 TTACCAATGGATCAA----- 899
Db |||:|||||
QY 1023 ValProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGluProGlyThr 1042
Db |||:|||||
QY 900 GAGTTTACTTACCGTGT-----AAAAATCGGGAACAAGCTTATAGGATC----- 944
Db |||:|||||
QY 1043 GluTyrThrIleTyrValIleAlaLeuLysAsnAsnGlnLysSerGluProLeuIleGly 1062
Db |||:|||||
QY 945 AATAAAAAATCTGGTCTGAATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATAT 1004
Db |||:|||||
QY 1063 ArgLysLysThrGly---GlnGluAlaLeuSerGlnThrIle----- 1076
Db |||:|||||
QY 1005 TAGCTCTTAAAAAAGGGGAAACCGTATGATCCCTTTGATCGCAGTCACCTTGAACACTG 1064
Db |||:|||||
QY 1077 -----SerTrpAlaProPheGln----- 1082
Db |||:|||||
QY 1065 TTCACCATCAATACGTTGATGTCGATACCAACGAATTCGTAATAAAGTGAGCAGCTCTTA 1124
Db |||:|||||
QY 1083 -----AspThrSerGluTyrIleIleSerCysHisProVal 1094
Db |||:|||||
QY 1125 ACAGTAGCGAAGTAACATTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAA 1184
Db |||:|||||
QY 1095 GlyThrAspGluGluProLeuGlnPheArg-----ValProGlyThrSerThrSer 1111
Db |||:|||||
QY 1185 CTACTCTACAACTCGATCGCTTTGTGTATATGAGACTATACCTTAACCTGGAAGTA 1244
Db |||:|||||
QY 1112 Ala-----ThrLeuThrGlyLeuThr 1118
Db |||:|||||
QY 1245 GAGGATAATCACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGAA 1304
Db |||:|||||
QY 1118 ----- 1118
Db |||:|||||
QY 1305 GGAGAGAAATGTAGTATCAT---TTAGCCGGTGGTGTGTCAGGCGCAGCAAAATGGTTTCAG 1361
Db |||:|||||
QY 1119 ---ArgGlyAlaThrTyrAsnIleIleValGluAlaLeuLysAspGlnGlnArgHisLys 1137
Db |||:|||||
QY 1362 CCCGAGTCCCGGTGCTGCTCAGTCAAGCAAGCCCGGTGTGTATGACAAATGGAACAC 1421
Db |||:|||||
QY 1138 ValArgGluGluValValThrValGlyAsnSerGlyTrpCysHisAspAsnGlyValAsn 1157
Db |||:|||||
QY 1422 TATCAGATAAATCAACAGTGGGAGCGGACC---TACCTAGGTAATGTGTGGTTGTACT 1478
Db |||:|||||
QY 1158 TyrLysIleGlyGluLysTrpAspArgGlnGlyGluAsnGlyGlnMetMetSerCysThr 1177
Db |||:|||||
QY 1479 TGTTATGAGGAAGCCGAGGT---TTTAACTCGAAAGTAAACCTGAAGCTGAAGAGACT 1535
Db |||:|||||
QY 1178 CysLeuGlyAsnGlyLysGlyGluPheLysCysAspProHis-----GluAlaThr 1194
Db |||:|||||
QY 1536 TGCTTTGACAGTACACTGGGAACACTTACGAGTGGGTGACACACTTATGAGCGTCTCTAAA 1595
Db |||:|||||
QY 1195 CysTyrAsp-----AspGlyLysThrTyrHisValGlyGluGlnTrpGlnLysGluTyr 1212
Db |||:|||||
QY 1596 GACTCCATGATCGGAGCTGTACCTGCTCGGCGGTGGCGGAGG 1640
Db |||:|||||
QY 1213 LeuGlyAlaIleCysSerCysThrCysPheGlyGlyGlnArgGly 1227
Db |||:|||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:35 ; Search time 25.4546 Seconds
(without alignments)
12556.953 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1: gcaacccgcagcctagcc.....gaataagctgtaccatctaa 1661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTC_spool_p/US09940235/runat_27012006_144218_27578/app_query.fasta_1.7708
-DB=PIR -QFWI=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptl -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN 1 1 185 @runat_27012006_144218_27578 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	440	1 A22801	streptokinase prec
2	1951.5	66.4	415	1 B280	streptokinase (EC
3	1942	66.1	440	2 S02723	streptokinase G pr
4	1779	60.5	440	2 S02724	streptokinase A pr
5	1673	56.9	414	2 JU0292	streptokinase - St
6	1670	56.8	440	2 S04168	streptokinase A pr
7	667	22.7	197	2 S53334	streptokinase - St
8	620	21.1	128	2 S77671	streptokinase A (E
9	600	20.4	2386	1 FNHU	fibronectin precu
10	575	19.6	2265	1 FNBO	fibronectin - bovi
11	574	19.5	2477	2 S14428	fibronectin precu
12	569	19.4	128	2 S77680	streptokinase A (E
13	557	19.0	128	2 S77688	streptokinase A (E
14	552	18.8	128	2 S77679	streptokinase A (E

15	496	16.9	2481	2 A43908	fibronectin - Afri
16	423	14.4	128	2 S77676	streptokinase A (E
17	419	14.3	128	2 S77673	streptokinase A (E
18	419	14.3	128	2 S77687	streptokinase A (E
19	417	14.2	128	2 S77682	streptokinase A (E
20	416	14.2	128	2 S77678	streptokinase A (E
21	408	13.9	128	2 S77677	streptokinase A (E
22	407	13.8	128	2 S77685	streptokinase A (E
23	401	13.6	128	2 S77683	streptokinase A (E
24	398	13.5	128	2 S77686	streptokinase A (E
25	397	13.5	128	2 S77684	streptokinase A (E
26	396	13.5	128	2 S77681	streptokinase A (E
27	390	13.3	128	2 S77674	streptokinase A (E
28	386	13.1	128	2 S77675	streptokinase A (E
29	171	5.8	1020	2 A29355	fibronectin - Chic
30	164	5.6	103	2 A49173	fibronectin - mous
31	147.5	5.0	2231	2 D71870	hypothetical prote
32	144	4.9	190	2 I51279	fibronectin - east
33	133.5	4.5	823	2 A90761	hypothetical prote
34	132	4.5	732	2 B90975	hypothetical prote
35	130.5	4.4	1315	2 T28679	fibrinogen-binding
36	130.5	4.4	1745	2 A46431	tight junction-ass
37	129	4.4	823	2 F85624	hypothetical prote
38	128.5	4.4	860	2 T39502	hypothetical prote
39	125	4.3	503	2 JQ1019	site-specific DNA-
40	125	4.3	823	2 B85822	hypothetical prote
41	125	4.3	4688	2 F82885	hypothetical prote
42	124.5	4.2	967	2 S25081	phosphoenolpyruvat
43	124.5	4.2	1237	2 AC1583	internalin protein
44	124	4.2	2334	2 S32920	cell wall-associat
45	123.5	4.2	953	2 C89824	hypothetical prote

ALIGNMENTS

RESULT 1
A22801
streptokinase precursor - Streptococcus "equisimilis"
C;Species: Streptococcus "equisimilis"
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A22801
R;Malke, H.; Roe, B.; Ferretti, J.J.
Gene 34, 357-362, 1995
A;Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H
A;Reference number: A22801; MUID:85232082; PMID:2989113
A;Accession: A22801
A;Molecule type: DNA
A;Residues: 1-440 <MAL>
A;Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:9407876; PIDN
A;Experimental source: strain H46A
C;Genetics:
A;Gene: skc
C;Superfamily: streptokinase

Alignment Scores:
Pred. No.: 3.95e-126 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.51% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x A22801 (1-440)

QY	183	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCATCTGTCAACAACAGCAATTCGTTGTT	242
Db	27	IlEaLaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal	46
QY	243	AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC	302
Db	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	66
QY	303	GATCTAACATCAGCAGCTGCTCATGGAGGAGAAAGACAGACAGCGCTTAAGTCCAAATCA	362

Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTCAGAAAGCTGACTACTA 422
 Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
 QY 423 AAGGCTATTCAAGAACAAATTGATCGGTAAACGTCACAGTCCACAGTAACGACGACTACTTTGAGGTC 482
 Db 107 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
 QY 483 ATTGATTTTGCAGCGATCAACCACTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 542
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 543 AAAGATGGTTCCGTTACCTGGCGACCCCAACCTGTCCAAGAATTTTGTCTAAACGGGACAT 602
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
 QY 603 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTG 662
 Db 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 663 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 722
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 723 GATACATAGCTTATGAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATACTA 782
 Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACTGAC 842
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCACATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 903 TTTACTACCGGTGTTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGCTGT 962
 Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 1022
 Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 1023 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTCAAACTGTTCAACCATCAATACGTT 1082
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1083 GATGTCGATACCAACGAATTGCTAAAGATGAGCAGCTCTTAAACAGCTACGCAACGTAAC 1142
 Db 327 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTATACCATCTCTGTGATAGGCTTAAACTACTCTACACAATCTC 1202
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1203 GATGCTTTGGTATTATGAGCATATACCTTAACTTGAAGAAAGTAGAGGATATACGATGAC 1262
 Db 367 AspAlaPheGlyIleLysAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1263 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCGAAGGAGAGAAATGCTAGCTAT 1322
 Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1323 CATTTAGCC 1331
 Db 407 HisLeuAla 409

RESULT 2
 BZSO

streptokinase (EC 3.4.-.-) - Streptococcus sp.
 C:Species: Streptococcus sp.
 C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
 C:Accession: A00967
 R:Jackson, K.W.; Tang, J.
 Biochemistry 21, 6620-6625, 1982
 A:Title: Complete amino acid sequence of streptokinase and its homology with serine protease
 A:Reference number: A00967; MUID:83127125; PMID:6760891
 A:Accession: A00967
 A:Molecule type: protein
 A:Residues: 1-415 <JAC>
 A:Cross-references: UNIPROT:P00779; UNIPARC:UPI0000172BF5
 A:Note: 169-Asp and 181-Asp were also found
 A:Note: this protein is not a protease, but it activates plasminogen by complexing with
 C:Superfamily: streptokinase
 C:Keywords: hydrolase

Alignment Scores:
 Pred. No.: 6,09e-124 Length: 415
 Score: 1951.50 Matches: 380
 Percent Similarity: 99.22% Conservative: 1
 Best Local Similarity: 98.96% Mismatches: 2
 Query Match: 66.40% Indels: 1
 DB: 1 Gaps: 1

US-09-940-235-10 (1-1661) x BZSO (1-415)

QY 183 ATTGCTGGAGCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATCGTGT 242
 Db 1 IleAlaGlyProGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATATTTTGAATC 302
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 303 GATCTAACATCAACGACCTGCTCATGAGAGAAAGACAGACGAAGCTTTAAGTCCAAAATCA 362
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 QY 363 AAACCATTTGCTACTGATGTCGCGGATGTCATAACTGAGAAAGCTGACTTACTA 422
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
 QY 423 AAGCTATTCAAGACAAATGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGTC 482
 Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 483 ATTGATTTTCAAGCGGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 542
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 543 AAAGATGGTTTCGGTAAACCTTTGCCGACCCCAACCTGTCTCAAGAAATTTTGTCTAAGCGGACAT 602
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 603 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 663 GAATATACCTGATGAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAACAGGCTCTCAA 722
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 723 GATACATAGCTATTGAAACCACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 782
 Db 181 LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAAACCAAGCGGCTATACGATTTATGAAACGTCAC 842
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 843 TCCTCAATCGTCACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902

Db	221	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	240
Qy	903	TTTACTACCGTGTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
Db	241	PheThrTyArgValLysAsnArgGluGlnAlaTyArgIleAsnLysLysSerGlyLeu	260
Qy	963	AATGAAGAATAAACAACTGACCTGATCTCTGAGAAATAT---TACGTCTTAAAAAA	1019
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerLeuGluTyLysTyValLeuLysLys	280
Qy	1020	GGGGAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAATAC	1079
Db	281	GlyGluLysProTyArgPheAspArgSerHisLeuLysLeuPheThrIleLysTy	300
Qy	1080	GTTGATGTCATACCAACGAATTTGCTAAAGTCAGCAGCTCTTAAACAGCTAGCGAACGT	1139
Db	301	ValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg	320
Qy	1140	AACCTAGACTTCAGAGATTTATACGATCTCTGCTGATAAGCTAAACTTCTTACAACAAT	1199
Db	321	AsnLeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLeuTyAsnAsn	340
Qy	1200	CTCGATGCTTTGGTATTATGACTATACCTTAACCTGCGAAAGTACAGGATATACCAT	1259
Db	341	LeuAspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAsp	360
Qy	1260	GACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGCAAGGAGAGATGCTAGC	1319
Db	361	AspThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSer	380
Qy	1320	TATCAATTTAGCC 1331	
Db	381	TyrHisLeuAla 384	
RESULT 3			
S02723			
streptokinase G precursor - Streptococcus sp. (strain 19908)			
C:Species: Streptococcus sp.			
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004			
C:Accession: S02723			
R:Walter, F.; Siegel, M.; Malke, H.			
Nucleic Acids Res. 17, 1262, 1989			
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.			
A:Reference number: S02723; MUID:89160265; PMID:2922269			
A:Accession: S02723			
A:Molecule type: DNA			
A:Residues: 1-440 <WAL>			
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:947095; PIDN			
C:Genetics:			
A:Gene: skg			
C:Superfamily: streptokinase			
F:1-26/Domain: signal sequence #status predicted <SIG>			
F:27-440/Product: streptokinase #status predicted <MAT>			
Alignment Scores:			
Pred. No.:	2,67e-123	Length:	440
Score:	1942.00	Matches:	375
Percent Similarity:	98.43%	Conservative:	2
Best Local Similarity:	97.91%	Mismatches:	6
Query Match:	66.08%	Indels:	0
DB:	2	Gaps:	0
US-09-940-235-10 (1-1661) x S02723 (1-440)			
Qy	183	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGT	242
Db	27	IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal	46
Qy	243	ACGGTGTCTGGTCTGTTGAGGGGAGCAATCAAGACATTTAGTCTTAAATTTTGTGAATC	302
Db	47	SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle	66
Qy	303	GATCTAACATCAGCACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA	362

Db	67	AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer	86
Qy	363	AAACCATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTTACTA	422
Db	87	LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106
Qy	423	AAGGCTATTCAAGAACAAATTCATCCCTAAACGTCACAGTAAACGACGACTTCTTTCAGGTC	482
Db	107	LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyPheGluVal	126
Qy	483	ATTGATTTTGAACGATGCAACCATTTACTGATGAAACCGCAAGGTCTACTTTGCTGCAC	542
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyPheAlaAsp	146
Qy	543	AAAGATGTTGGTAAACCTTCCGCCACCACTGCTCCCAAGATTTTTCCTAAGCGGACAT	602
Db	147	LysAspGlySerValThrLeuProIleGlnProValGlnGluPheLeuLeuLysGlyHis	166
Qy	603	GTGCGCGTTAGACCATATAAGAAAAACAATACAAAAACCAAGCGAAATCTGTTGATGTG	662
Db	167	ValArgValArgProTyLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186
Qy	663	GAATATATCTGTACAGTTTACTCCCTTAAACCTCTGATGACGATTTTCAGACCGAGTCTCAA	722
Db	187	GluTyThrValGlnPheThrProLeuAsnProAspAspAspPheArgProAlaLeuLys	206
Qy	723	GATCTAAGCTATTGAAAAACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA	782
Db	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	226
Qy	783	GCTCAAGCAACAAAGCATTTTAAACAAACCAACCAACGAGCTATACGATTTTGAACGTGAC	842
Db	227	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGluArgAsp	246
Qy	843	TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCATCGATCAACAG	902
Db	247	SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu	266
Qy	903	TTTACTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
Db	267	PheThrTyHisValLysAsnArgGluGlnAlaTyArgIleAsnLysLysSerGlyLeu	286
Qy	963	AATCAAGAAATAACACACACTGACCTGATCTCTGAGAAATATTAGTCTTAAAAAAGGG	1022
Db	287	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyTyTyValLeuLysLysGly	306
Qy	1023	GAAGAAGCGTATGATCCCTTTGATGCGAGTCACTTTGAAACTGTTCAACCATCAAAATACGT	1082
Db	307	GluLysProTyArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal	326
Qy	1083	GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	1142
Db	327	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346
Qy	1143	TTAGACTTCAGAGATTTTATAGGATCCCTCGTGATAGGCTAAACTCTCTACAACAATCTC	1202
Db	347	LeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLeuTyAsnAsnLeu	366
Qy	1203	GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC	1262
Db	367	AspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspAsp	386
Qy	1263	ACCAACCGTATCAFAACCGTTTATATGGCAAGCGACCGCAAGGAGAGATGCTAGCTAT	1322
Db	387	ThrAsnArgIleIleThrValTyMetGlyLysArgProGluGlyGluAsnAlaSerTy	406
Qy	1323	CATTAGCC 1331	
Db	407	HisLeuAla 409	
RESULT 4			
S02724			

QY	Db	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6	Seq7	Seq8	Seq9	Seq10	Seq11	Seq12	Seq13	Seq14	Seq15	Seq16	Seq17	Seq18	Seq19	Seq20	Seq21	Seq22	Seq23	Seq24	Seq25	Seq26	Seq27	Seq28	Seq29	Seq30	Seq31	Seq32	Seq33	Seq34	Seq35	Seq36	Seq37	Seq38	Seq39	Seq40	Seq41	Seq42	Seq43	Seq44	Seq45	Seq46	Seq47	Seq48	Seq49	Seq50	Seq51	Seq52	Seq53	Seq54	Seq55	Seq56	Seq57	Seq58	Seq59	Seq60	Seq61	Seq62	Seq63	Seq64	Seq65	Seq66	Seq67	Seq68	Seq69	Seq70	Seq71	Seq72	Seq73	Seq74	Seq75	Seq76	Seq77	Seq78	Seq79	Seq80	Seq81	Seq82	Seq83	Seq84	Seq85	Seq86	Seq87	Seq88	Seq89	Seq90	Seq91	Seq92	Seq93	Seq94	Seq95	Seq96	Seq97	Seq98	Seq99	Seq100	Seq101	Seq102	Seq103	Seq104	Seq105	Seq106	Seq107	Seq108	Seq109	Seq110	Seq111	Seq112	Seq113	Seq114	Seq115	Seq116	Seq117	Seq118	Seq119	Seq120	Seq121	Seq122	Seq123	Seq124	Seq125	Seq126	Seq127	Seq128	Seq129	Seq130	Seq131	Seq132	Seq133	Seq134	Seq135	Seq136	Seq137	Seq138	Seq139	Seq140	Seq141	Seq142	Seq143	Seq144	Seq145	Seq146	Seq147	Seq148	Seq149	Seq150	Seq151	Seq152	Seq153	Seq154	Seq155	Seq156	Seq157	Seq158	Seq159	Seq160	Seq161	Seq162	Seq163	Seq164	Seq165	Seq166	Seq167	Seq168	Seq169	Seq170	Seq171	Seq172	Seq173	Seq174	Seq175	Seq176	Seq177	Seq178	Seq179	Seq180	Seq181	Seq182	Seq183	Seq184	Seq185	Seq186	Seq187	Seq188	Seq189	Seq190	Seq191	Seq192	Seq193	Seq194	Seq195	Seq196	Seq197	Seq198	Seq199	Seq200	Seq201	Seq202	Seq203	Seq204	Seq205	Seq206	Seq207	Seq208	Seq209	Seq210	Seq211	Seq212	Seq213	Seq214	Seq215	Seq216	Seq217	Seq218	Seq219	Seq220	Seq221	Seq222	Seq223	Seq224	Seq225	Seq226	Seq227	Seq228	Seq229	Seq230	Seq231	Seq232	Seq233	Seq234	Seq235	Seq236	Seq237	Seq238	Seq239	Seq240	Seq241	Seq242	Seq243	Seq244	Seq245	Seq246	Seq247	Seq248	Seq249	Seq250	Seq251	Seq252	Seq253	Seq254	Seq255	Seq256	Seq257	Seq258	Seq259	Seq260	Seq261	Seq262	Seq263	Seq264	Seq265	Seq266	Seq267	Seq268	Seq269	Seq270	Seq271	Seq272	Seq273	Seq274	Seq275	Seq276	Seq277	Seq278	Seq279	Seq280	Seq281	Seq282	Seq283	Seq284	Seq285	Seq286	Seq287	Seq288	Seq289	Seq290	Seq291	Seq292	Seq293	Seq294	Seq295	Seq296	Seq297	Seq298	Seq299	Seq300	Seq301	Seq302	Seq303	Seq304	Seq305	Seq306	Seq307	Seq308	Seq309	Seq310	Seq311	Seq312	Seq313	Seq314	Seq315	Seq316	Seq317	Seq318	Seq319	Seq320	Seq321	Seq322	Seq323	Seq324	Seq325	Seq326	Seq327	Seq328	Seq329	Seq330	Seq331	Seq332	Seq333	Seq334	Seq335	Seq336	Seq337	Seq338	Seq339	Seq340	Seq341	Seq342	Seq343	Seq344	Seq345	Seq346	Seq347	Seq348	Seq349	Seq350	Seq351	Seq352	Seq353	Seq354	Seq355	Seq356	Seq357	Seq358	Seq359	Seq360	Seq361	Seq362	Seq363	Seq364	Seq365	Seq366	Seq367	Seq368	Seq369	Seq370	Seq371	Seq372	Seq373	Seq374	Seq375	Seq376	Seq377	Seq378	Seq379	Seq380	Seq381	Seq382	Seq383	Seq384	Seq385	Seq386	Seq387	Seq388	Seq389	Seq390	Seq391	Seq392	Seq393	Seq394	Seq395	Seq396	Seq397	Seq398	Seq399	Seq400	Seq401	Seq402	Seq403	Seq404	Seq405	Seq406	Seq407	Seq408	Seq409	Seq410	Seq411	Seq412	Seq413	Seq414	Seq415	Seq416	Seq417	Seq418	Seq419	Seq420	Seq421	Seq422	Seq423	Seq424	Seq425	Seq426	Seq427	Seq428	Seq429	Seq430	Seq431	Seq432	Seq433	Seq434	Seq435	Seq436	Seq437	Seq438	Seq439	Seq440	Seq441	Seq442	Seq443	Seq444	Seq445	Seq446	Seq447	Seq448	Seq449	Seq450	Seq451	Seq452	Seq453	Seq454	Seq455	Seq456	Seq457	Seq458	Seq459	Seq460	Seq461	Seq462	Seq463	Seq464	Seq
----	----	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	-----

Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 363 AAACCAATTTGCTACTGATAGTGGCGCATGTGCATAAACCTTGAGAAAGCTGACTACTA 422
Db 61 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 423 AAGGCTATTCAAGAACAAATGATCGCTAAAGCTGCACAGTAAACGACGACTACTTTGAGTGC 482
Db 81 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 100
QY 483 ATTGATTTTGCAGCGCATGCAACCATCTACTGATCGAAACGGCAGGCTTCTGCTGAC 542
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 543 AAAGATGTTTCCGCTAACTTGGCCGACCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 602
Db 121 ArgAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
QY 603 GTGCGGCTTAGACATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
Db 141 ValArgValArgProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 160
QY 663 GAATATACGTACAGATTACTCCCTTAAACCCCTGATGACGATTCAGACCGAGTCTCAA 722
Db 161 AsnTyrGluValSerPheValSerGluThrGlyAsnLeuAspPheThrProSerLeuLys 180
QY 723 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAACTACTA 782
Db 181 GluArgTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerGlnGluLeuAla 200
QY 783 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTATGAAACGTGAC 842
Db 201 AlaIleAlaGlnPheIleLeuSerLysGluHisProAspTyrIleIleThrLysArgAsp 220
QY 843 TCTCTAATCGTCTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 903 TTACTTACCGTCTTAAATCGGACACGCTTATAGGATCAATAAAATCTGGTCTG 962
Db 241 PheThrTyrHisIleLysAspArgGlnGlnAlaTyrLysAlaAsnSerLysThrGlyIle 260
QY 963 AATGAAGAATAAACCACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1022
Db 261 ValGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCACCATCAATACGTT 1082
Db 281 GluGluProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1083 GATGTCGATACCAACCAATGCTTAAAGTGACGACGCTCTTAACAGCTAGCAAGCTAAC 1142
Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1143 TTAGACTTCAGAGATTATACGATCTCTGCTGATAAGGCTAACTCTACACAATCTC 1202
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACTTGGAAAGTAGAGGATAATCACGATGAC 1262
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsnAsp 360
QY 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGAGAGAAATGCTAGCTAT 1322
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1323 CATTTAGCC 1331
Db 381 HisLeuAla 383
RESULT 6
S04168
streptokinase A precursor - Streptococcus pyogenes (strain NZ131)

C:Species: Streptococcus pyogenes
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: S04168
R:Huang, T.T.; Malke, H.; Ferretti, J.J.
Mol. Microbiol. 3, 197-205, 1989
A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri-
A:Reference number: S04168; MUID:89343623; PMID:2668686
A:Accession: S04168
A:Molecule type: DNA
A:Residues: 1-440 <HUA>
A:Cross-references: UNIPROT:Q57391; UNIPARC:UPI00000BD04A; EMBL:X51517; NID:g47437; PID:
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-440/Product: streptokinase A #status predicted <MAT>
Alignment Scores:
Pred. NO.: 5.68e-105 Length: 440
Score: 1670.00 Matches: 324
Percent Similarity: 90.34% Conserved: 22
Best Local Similarity: 84.60% Mismatches: 37
Query Match: 56.82% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-10 (1-1661) x S04168 (1-440)

QY 183 ATTCTGACCTGAGTGGCTGCTAGACGCTCCATCTGCTCAACACAGCAATGGTTGTT 242
Db 27 IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 243 AGCGTGTCTGCTACTGTTGAGGGCAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 302
Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnGluIleSerLeuLysPheGluIle 66
QY 303 GATCTAATCATCAGCACTGCTCATGAGAGAAAGACAGACGCAAGGCTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACATTTGCTACTGATAGTGGCGCATGTGCATAACTTGAGAAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAspLysGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 106
QY 423 AAGCTATTCAAGAACAAATGATCGCTAAACGTCCACATACGACGACTACTTTGAGTGC 482
Db 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 483 ATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 543 AAAGATGTTTGGTAACTTTGCCGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT 602
Db 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 603 GTGCGGCTTAGACATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 662
Db 167 ValArgValLysProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 186
QY 663 GAATATACGTCTGACATTTTACTCCCTTAAACCCCTGATGACGATTCAGACCGGCTCAA 722
Db 187 AsnTyrGluValSerPheValSerGluThrGlyAspLeuAspPheThrProLeuLeu 206
QY 723 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 782
Db 207 AsnGlnTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerGlnGluLeuAla 226
QY 783 GCTCAAGCAACAAAGCATTTTAAACAAAAACCAAGGCTATACGATTTATGAAACGTGAC 842
Db 227 AlaIleAlaGlnPheIleLeuSerLysLysHisProAspTyrIleIleThrLysArgAsp 246
QY 843 TCCTCAATCTCTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902

Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy 903 TTTACTACCGTGTAAATAACCGGAACAAGCTTATAGGATCAATAAAAAATCTGCTCG 962
Db 267 PheThrThrHisIleLysAspArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 286
Qy 963 AATGAAGAAATAACAACACTGACCTGATCTCTCGAATAATATTACGTCCTTAAAAAGGG 1022
Db 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
Qy 1023 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTGAAACTGTGTCAACATCAAAATACGTT 1082
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
Qy 1083 GATGTCGATACCAACGAATTGCTTAAAGTGAAGCTCTTAACAGCTAGCGCAACCTAAC 1142
Db 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Qy 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTTAAACTACTCTACACAATCTC 1202
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Qy 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1262
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
Qy 1263 ACCAACCCTATACATACCGTTTATATGGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1322
Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
Qy 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 7
S53334
streptokinase - Streptococcus sp.
C;Species: Streptococcus sp.
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S53334
R;Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
Biochem. J. 304, 235-241, 1994
A;Title: Function of streptokinase fragments in plasminogen activation.
A;Reference number: S53334; MUID:95091634; PMID:7998939
A;Accession: S53334
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-197 <SH1>
A;Cross-references: UNIPROT:Q7M115; UNIPARC:UPI000017AC2C

Alignment Scores:
Pred. No.: 2,09e-37 Length: 197
Score: 667.00 Matches: 167
Percent Similarity: 44.71% Conservative: 2
Best Local Similarity: 44.18% Mismatches: 1
Query Match: 22.69% Indels: 208
DB: 2 Gaps: 9

US-09-940-235-10 (1-1661) x S53334 (1-197)

Qy 183 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGT 242
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsn- 14
Qy 243 AGCGTGTGCTGCTACTGTTGAGGGGCGAATCAAGACATAGTCTTAAATTTTGAATC 302
Db 14 ----- 14
Qy 303 GATCTAACATACACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCAAATCA 362
Db 15 -----Ser 15
Qy 363 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 422

Db 16 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 35
Qy 423 AGGCTATTCAAGAACAAATTGATCGTAACTCCACAGTAAACGACGACTACTTTGAGTC 482
Db 36 LysAlaIleGlnTrpGlnIleLeu- 43
Qy 483 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGTCTGAC 542
Db 44 -----AsnGlyLysValTyrPheAlaAsp 51
Qy 543 AAAGATGGTTCGGTAACCTTTCGCGACCACTGTCCAAGAATTTTGTCTAAGCGGCAT 602
Db 52 LysAspGlySerValThr- 57
Qy 603 GTGCGGTTAGACCATTATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 58 -----GluLysProIleGlnAsnGlnAlaLysSerValAspVal 70
Qy 663 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAA 722
Db 71 Glu- 71
Qy 723 GATCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 71 ----- 71
Qy 783 GCTCAAGCACAAGCAATTTTAAACAAACACCCAGGCTATACGATTTATGATGAACGTGAC 842
Db 72 -----AsnHisProGlyTyrThrIleTyrGluArgAsp 82
Qy 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Db 83 SerSerIleVal- 94
Qy 903 TTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTTGCTG 962
Db 95 PheThrTyr- 100
Qy 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
Db 101 AsnGluGluIleAsnAsnThrAspLeuIleSer- 111
Qy 1023 GAAAGCCGTATGATCCCTTTGATCGCAGCTCACTTGAAACTGTTCCACCAATACGTT 1082
Db 112 -----TyrVal 113
Qy 1083 GATGTCGATACCAACCAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAAC 1142
Db 114 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 133
Qy 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATTAAGGCTAAACTACTCTACACAATCTC 1202
Db 134 -----AspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 149
Qy 1203 GATGCTTTTGGTATTATGACTTATACCTTAACCTGGAAGAGTAGAGGATATACGATGAC 1262
Db 150 AspAlaPheGlyIleMet- 155
Qy 1263 ACCAACCCTGATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGAAATGCT 1316
Db 156 -----IleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAla 170

RESULT 8
S77671
streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
C;Species: Streptococcus pyogenes
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
C;Accession: S77671; S77672
R;Kapoor, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A;Reference number: S77671; MUID:96037795; PMID:7565111

A;Residues: 1948-2067 <UMB>
A;Cross-references: UNIPARC:UPI000000046A; GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:
A;Accession: B32245
A;Molecule type: mRNA
A;Residues: 1975-1991,2017-2039 <UM2>
A;Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I5273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2113-2127 <SEK>
A;Cross-references: UNIPARC:UPI00006E04C; GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <K03>
A;Cross-references: UNIPARC:UPI0000174334; GB:K00799; NID:g182681; PIDN:AAA52460.1; PID:
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GAR1>
A;Cross-references: UNIPARC:UPI0000174335
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the R
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300,551-560 <GAR2>
A;Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GRI>
A;Cross-references: UNIPARC:UPI0000174338
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'Q',679-703,'PT' <CAL>
A;Cross-references: UNIPARC:UPI0000174339
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Cross-references: UNIPARC:UPI0000141CD5
A;Note: residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2058 <GAR3>
A;Cross-references: UNIPARC:UPI000017433A
R;Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand

Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
A;Cross-references: UNIPARC:UPI000017433B
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-da
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080;2112-2356 <GAR4>
A;Cross-references: UNIPARC:UPI000017433C; UNIPARC:UPI000017433D
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins
ation, and transformation.
C;Genetics:
A;Gene: GDB:FNI
A;Cross-references: GDB:119135; OMIM:135600
A;Map position: 2q34-2q34
A;Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology;
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplica
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <FHB>
F;52-87/Domain: fibronectin type I repeat homology <1F1>
F;97-135/Domain: fibronectin type I repeat homology <1F2>
F;141-179/Domain: fibronectin type I repeat homology <1F3>
F;186-225/Domain: fibronectin type I repeat homology <1F4>
F;231-270/Domain: fibronectin type I repeat homology <1F5>
F;308-608/Domain: collagen binding <CBR>
F;360-401/Domain: fibronectin type II repeat homology <1F6>
F;420-461/Domain: fibronectin type II repeat homology <2F1>
F;470-508/Domain: fibronectin type II repeat homology <2F2>
F;518-555/Domain: fibronectin type I repeat homology <1F8>
F;561-599/Domain: fibronectin type I repeat homology <1F9>
Alignment Scores:
Pred. No.: 7-72e-33 Length: 2386
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.42% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-10 (1-1661) x FNHU (1-2386)
QY 1341 CAGGCGCAGCAATGGTTTCAGCCCGTCCCGTGGCTCTCAGTCAAGCAAGCCCGGT 1400
Db 32 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 51
QY 1401 TGTATGACATCGAAGAAACACTATCAGATAAATCAACAGTGGGCGGACCTACTAGGT 1460
Db 52 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 71
QY 1461 AATGTGTGTGGTTGTTACTTGTATGAGGAAGCCGAGGTGTTTAACTCGGAAAGTAAACCT 1520
Db 72 AsnValLeuValCysThrCysTyrGlyArgGlyPheAsnGlyGluSerLysPro 91
QY 1521 GAAGCTGAACAGACTTGTCTTTTGACAGTACACTGGGAACACTTACCAGGTGGGTGCACACT 1580
Db 92 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 111
QY 1581 TATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTGACCTGCGATCGGGGCTGGCGGAGG 1640
Db 112 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 131
QY 1641 AGAATAAGCTGTACCATC 1658

Db ||||| 132 ArgileSerCysThrille 137

RESULT 10

FNBO

fibronectin - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

A:Accession: A26452; A21165; A23292

R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A:Title: Complete primary structure of bovine plasma fibronectin.

A:Reference number: A26452; MUID:87054047; PMID:3780752

A:Accession: A26452

A:Molecule type: protein

A:Residues: 1-2265 <SKO>

A:Cross-references: UNIPROT:P07589; UNIPARC:UPI0000012A7BE

R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983

A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

A:Reference number: A21165; MUID:83221567; PMID:6304699

A:Accession: B21165

A:Molecule type: mRNA

A:Residues: 2170-2265 <KOR>

A:Cross-references: UNIPARC:UPI000016C30C; GB:K00800; NID:G163055; PIDN:AAA30521.2; PID:R:Pedersen, T.E.; Thogersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottor Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983

A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal

A:Reference number: A23292; MUID:83117805; PMID:6218503

A:Accession: A23292

A:Molecule type: protein

A:Residues: 1-16, 'C', 18-20, 'S', 22-432; 447-463; 1367-1517; 1567-1673; 2062-2176, 'N', 2178-226

A:Cross-references: UNIPARC:UPI000017433B; UNIPARC:UPI000017433F; UNIPARC:UPI0000174340;

C:Comment: Cys-1201 and Cys-201 have free sulphydryl groups.

C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b

C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib

C:Comment: aling, and maintenance of cell shape.

C:Comment: Plasma fibronectin is synthesized by hepatocytes.

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe

C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu

F:21-241/Domain: fibrin and heparin binding <FBR>

F:21-56/Domain: fibronectin type I repeat homology <1F1>

F:166-104/Domain: fibronectin type I repeat homology <1F2>

F:110-148/Domain: fibronectin type I repeat homology <1F3>

F:155-194/Domain: fibronectin type I repeat homology <1F4>

F:200-239/Domain: fibronectin type I repeat homology <1F5>

F:277-577/Domain: collagen binding <CBR>

F:277-311/Domain: fibronectin type II repeat homology <1F6>

F:329-370/Domain: fibronectin type II repeat homology <2F1>

F:389-430/Domain: fibronectin type II repeat homology <2F2>

F:439-477/Domain: fibronectin type I repeat homology <1F7>

F:487-524/Domain: fibronectin type I repeat homology <1F8>

F:530-568/Domain: fibronectin type I repeat homology <1F9>

F:578-661/Domain: fibronectin type III repeat homology <FN3A>

F:688-770/Domain: fibronectin type III repeat homology <FN3B>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

F:875-957/Domain: fibronectin type III repeat homology <FN3D>

F:965-1046/Domain: fibronectin type III repeat homology <FN3E>

F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>

F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>

F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>

F:1326-1404/Domain: fibronectin type III repeat homology <GN3I>

F:1410-1517/Domain: cell attachment <CAD>

F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>

F:1493-1495/Region: cell attachment (R-G-D) motif

F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>

F:1600-1870/Domain: heparin binding <HB2>

F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>

F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>

F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>

F:1970-1972/Region: cell attachment (R-G-D) motif

F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>

F:1985-2216/Domain: fibrin binding <FB2>

F:2085-2124/Domain: fibronectin type I repeat homology <1F10>

F:2130-2167/Domain: fibronectin type I repeat homology <1F11>

F:2174-2209/Domain: fibronectin type I repeat homology <1F12>

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Cross-link: isopeptide (Gln) (interchain to lys N6-amino of fibrin) #status experimental

F:21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-7,2155-2167,2174-2200,2198-2209/Diulfide bonds: #status predicted

F:399,497,511,846,976,1213,1987/Binding site: carboxylate (Asn) (covalent) #status absent

F:1205,1692/Binding site: carboxylate (Asn) (covalent) #status absent

F:1943,1944/Binding site: carboxylate (Thr) (covalent) #status experimental

F:2246/Disulfide bonds: interchain (to 2250) #status predicted

F:2250/Disulfide bonds: interchain (to 2246) #status predicted

F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Alignment Scores:

Pred. No.:	3,73e-31	Length:	2265
Score:	575.00	Matches:	100
Percent Similarity:	97.17%	Conservative:	3
Best Local Similarity:	94.34%	Mismatches:	3
Query Match:	19.56%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-10 (1-1661) x FNBO (1-2265)

Qy 1341 CAGCGCAGCAATGGTTTCAGCCCGGTCGCTGCTCAGTCAAGCAAGCCCGGT 1400

Db 1 GlnAlaGlnGlnIleValGlnProGlnSerProLeuThrValSerGlnSerIysProGly 20

Qy 1401 TGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 1460

Db 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40

Qy 1461 AATGTGTGGTTCTACTTGTATGAGGAGCGGAGCGGAGGTTTAACTGCGAAGTAAACCT 1520

Db 41 SerAlaLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerIysPro 60

Qy 1521 GNACTCAAGACACTGCTGTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 1580

Db 61 GluProGluGluThrCysPheAspIysTyrThrGlyAsnThrTyrArgValGlyAspThr 80

Qy 1581 TATGAGCGTCTTAAAGACTCATGATCTGGGACTGTACTGCATCGGGGCTGGCGGAGGG 1640

Db 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100

Qy 1641 AGAATAAGCTGTACATC 1658

Db 101 ArgIleSerCysThrIle 106

RESULT 11

S14428

fibronectin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004

C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049

R:Hynes, R.O.

submitted to the EMBL Data Library, July 1989

A:Reference number: S14428

A:Accession: S14428

A:Molecule type: mRNA

A:Residues: 1-2477 <HYN>

A:Cross-references: UNIPROT:P04937; UNIPARC:UPI000012A7C6; EMBL:X15906; NID:G56163; PID:R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.

EMBO J. 6, 2573-2580, 1987

A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.

A:Reference number: S12455; MUID:88054951; PMID:2445560

A:Accession: S12455

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 609-1810, 'T', 1812-2283 <SCH>

A:Cross-references: UNIPARC:UPI0000177A6A; EMBL:X15906

R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.

Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984

A:Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing

A;Reference number: A22319; MUID:84298097; PMID:6089177

A;Accession: A22319

A;Molecule type: DNA

A;Residues: 2032-2237 <TAM>

A;Cross-references: UNIPARC:UPI0000177ARB

R;Falkenberg, C.; Enghild, J.J.; Thogersen, I.B.; Salvesen, G.; Akerstroem, B.

Biochem. J. 301, 745-751, 1994

A;Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in

A;Reference number: S46203; MUID:94330948; PMID:7519849

A;Accession: S46203

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1183-1192; 'GLN', 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PY', 1385-1399 <PAL>

A;Cross-references: UNIPARC:UPI0000177ARC; UNIPARC:UPI0000177AED; UNIPARC:UPI0000177AEE

R;Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.

EMBO J. 6, 2565-2572, 1987

A;Title: Organization of the fibronectin gene provides evidence for exon shuffling during

A;Reference number: S00459; MUID:88054950; PMID:3119323

A;Accession: S00459

A;Molecule type: DNA

A;Residues: 1-139;2382-2477 <PAT>

A;Cross-references: UNIPARC:UPI000017095C; UNIPARC:UPI0000177ABF; EMBL:X05831

A;Note: the authors translated the codon CCT for residues 51 and 94 as Ala

R;Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.

Cell 35, 421-431, 1983

A;Title: Three different fibronectin mRNAs arise by alternative splicing within the codi

A;Reference number: A27252; MUID:84082067; PMID:6317187

A;Accession: A27252

A;Molecule type: mRNA

A;Residues: 1586-1720, 'T', 1722, 1813-2477 <SC2>

A;Cross-references: UNIPARC:UPI0000177AF0

R;Odermatt, E.; Tamkun, J.W.; Hynes, R.O.

Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985

A;Title: Repeating modular structure of the fibronectin gene: Relationship to protein su

A;Reference number: 159049; MUID:86016741; PMID:3863113

A;Accession: 159049

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1722-1810 <RRS>

A;Cross-references: UNIPARC:UPI000000040C; GB:M11750; NID:g204164; PIDN:AAA41170.1; PID:

C;Genetics:

A;Introns: 51/1; 94/1; 2416/3; 2454/3

C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe

C;Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; dupli

F;1-32/Domain: signal sequence #status predicted <SIG>

F;1-32/2477/Product: fibronectin #status predicted <MAT>

F;53-88/Domain: fibronectin type I repeat homology <F1>

F;98-136/Domain: fibronectin type I repeat homology <F2>

F;142-180/Domain: fibronectin type I repeat homology <F3>

F;187-226/Domain: fibronectin type I repeat homology <F4>

F;232-271/Domain: fibronectin type I repeat homology <F5>

F;308-342/Domain: fibronectin type I repeat homology <F6>

F;360-401/Domain: fibronectin type II repeat homology <F1>

F;420-461/Domain: fibronectin type II repeat homology <F2>

F;470-508/Domain: fibronectin type I repeat homology <F7>

F;518-555/Domain: fibronectin type I repeat homology <F8>

F;561-599/Domain: fibronectin type I repeat homology <F9>

F;609-692/Domain: fibronectin type III repeat homology <FN3A>

F;718-800/Domain: fibronectin type III repeat homology <FN3B>

F;809-890/Domain: fibronectin type III repeat homology <FN3C>

F;905-987/Domain: fibronectin type III repeat homology <FN3D>

F;995-1076/Domain: fibronectin type III repeat homology <FN3E>

F;1085-1164/Domain: fibronectin type III repeat homology <FN3F>

F;1172-1257/Domain: fibronectin type III repeat homology <FN3G>

F;1265-1348/Domain: fibronectin type III repeat homology <FN3H>

F;1356-1439/Domain: fibronectin type III repeat homology <FN3I>

F;1447-1529/Domain: fibronectin type III repeat homology <FN3J>

F;1537-1619/Domain: fibronectin type III repeat homology <FN3K>

F;1614-1616/Region: cell attachment (R-G-D) motif

F;1631-1713/Domain: fibronectin type III repeat homology <FN3L>

F;1721-1803/Domain: fibronectin type III repeat homology <FN3M>

F;1811-1893/Domain: fibronectin type III repeat homology <FN3N>

F;1903-1984/Domain: fibronectin type III repeat homology <FN3O>

F;1992-2074/Domain: fibronectin type III repeat homology <FN3P>

F;2181-2183/Region: cell attachment (R-G-D) motif

F;2193-2273/Domain: fibronectin type III repeat homology <FN3Q>

F;2296-2335/Domain: fibronectin type I repeat homology <F10>

F;2341-2378/Domain: fibronectin type I repeat homology <F11>

F;2385-2420/Domain: fibronectin type I repeat homology <F12>

F;53-79, 77-88, 98-126, 124-136, 142-170, 168-180, 187-216, 214-226, 232-261, 259-271, 308-335, 333

368, 2366-2378, 2385-2411, 2409-2420/Disulfide bonds: #status predicted

F;2458/Disulfide bonds: interchain (to 2462) #status predicted

F;2462/Disulfide bonds: interchain (to 2458) #status predicted

Alignment Scores:

Pred. No.:	4.37e-31	Length:	2477
Score:	574.00	Matches:	100
Percent Similarity:	97.17%	Conservative:	3
Best Local Similarity:	94.34%	Mismatches:	3
Query Match:	19.53%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-10 (1-1661) x S14428 (1-2477)

QY 1341 CAGCGCAGCAATGGTTTCAGCCCGTCCCGTGGCTGTCAGTCAAGCAAGCCCGT 1400

Db 33 GlnAlaGlnGlnIleValGlnProSerProValAlaValSerGlnSerLysProgly 52

QY 1401 TGTATGCAATCGAAAACATATCAGATAAATCAACAGTGGGAGCGACCTACTAGGT 1460

Db 53 CysPheAspAenGlyLysHisTyrGlnIleAenGlnInTrpGluArgThrTyrLeugly 72

QY 1461 AATGTGTGTGTTGTTACTTGTATTGAGGAGCGGAGGCTTTAACTGCGAAAGTAAACCT 1520

Db 73 AsnAlaLeuValCysThrCysTyrGlyGlySerArglyPheAenCysGluSerLysPro 92

QY 1521 GAAGCTGAAGAGACTTGTCTTTTGACAGTACACTCGGGAACACTTACCGAGTGGTGCACT 1580

Db 93 GluProGluGluThrCysPheAspLysTyrThrGlyAenThrTyrLysValGlyAspThr 112

QY 1581 TATGAGGCTCTTAAAGACTCATGATCTGGGACTGTACTGCATCGGGGCTGGGCGAGG 1640

Db 113 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArggly 132

QY 1641 AGAATAGCTGTACCATC 1658

Db 133 ArgIleSerCysThrIle 138

RESULT 12

S77680

C;Species: Streptococcus pyogenes

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C;Accession: S77680

R;Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser

Mol. Microbiol. 16, 509-519, 1995

A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A;Reference number: S77671; MUID:96037795; PMID:7565111

A;Accession: S77680

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAP>

A;Cross-references: UNIPROT:Q54687; UNIPARC:UPI00000BD2D3; EMBL:U25862; NID:g818926; PI:

A;Experimental source: strain ET51/ML7

C;Genetics:

A;Gene: ska

C;Superfamily: streptokinase

C;Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	8.23e-31	Length:	128
Score:	569.00	Matches:	110
Percent Similarity:	91.41%	Conservative:	7
Best Local Similarity:	85.94%	Mismatches:	11
Query Match:	19.36%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-10 (1-1661) x S77680 (1-128)

```

Qy 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATGTG 662
Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20
Qy 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCAAA 722
Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 40
Qy 723 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 41 AsnThrLysLeuLeuLysThrLeuAlaIleGlyThrValThrSerGlnGluLeu 60
Qy 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCTATACGATTATGAACGTGAC 842
Db 61 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 80
Qy 843 TCCTCAATCGTCACTCATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
Qy 903 TTTACTTACCGTGTAAATAACCGGAACAAGCTTTATAGGATCAATAAAAAATCTGCTGTG 962
Db 101 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysSerGlyGln 120
Qy 963 AATGAAGAATAAACCAACACTGAC 986
Db 121 GluGluLysThrAsnAsnThrAsp 128

```

RESULT 13

S77680
streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77688
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54695; UNIPARC:UPI00000BDBDC; EMBL:U25870; NID:g818942; PIR
A:Experimental source: strain ET76/M72
C:Genetics:
A:Gene: aka
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 5,29e-30 Length: 128
Score: 557.00 Matches: 107
Percent Similarity: 89.84% Conservative: 8
Best Local Similarity: 83.59% Mismatches: 13
Query Match: 18.95% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x S77688 (1-128)

```

Qy 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATGTG 662
Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20
Qy 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCAAA 722
Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 40
Qy 723 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 41 AsnThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60

```

```

Qy 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCTATACGATTATGAACGTGAC 842
Db 61 AlaGlnAlaGlnSerIleLeuIleLysHisProAspTyrThrIleTyrGluArgAsp 80
Qy 843 TCCTCAATCGTCACTCATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100
Qy 903 TTTACTTACCGTGTAAATAACCGGAACAAGCTTTATAGGATCAATAAAAAATCTGCTGTG 962
Db 101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspIle 120
Qy 963 AATGAAGAATAAACCAACACTGAC 986
Db 121 LysGluLysThrAsnAsnThrAsp 128

```

RESULT 14

S77679
streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77679
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77679
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54686; UNIPARC:UPI00000B8D6F; EMBL:U25861; NID:g818924; PIR
A:Experimental source: strain ET50/M43
C:Genetics:
A:Gene: aka
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 1.15e-29 Length: 128
Score: 552.00 Matches: 106
Percent Similarity: 89.06% Conservative: 8
Best Local Similarity: 82.81% Mismatches: 14
Query Match: 18.78% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x S77679 (1-128)

```

Qy 603 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATGTG 662
Db 1 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 20
Qy 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCAAA 722
Db 21 ArgTyrAlaValGlnPheThrProLeuAsnProAspAspPheThrProValLeuLys 40
Qy 723 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 782
Db 41 AsnThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60
Qy 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGCTATACGATTATGAACGTGAC 842
Db 61 AlaGlnAlaGlnSerIleLeuIleLysHisProAsnTyrThrIleHisGluArgAsp 80
Qy 843 TCCTCAATCGTCACTCATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100
Qy 903 TTTACTTACCGTGTAAATAACCGGAACAAGCTTTATAGGATCAATAAAAAATCTGCTGTG 962
Db 101 PheThrTyrHisValLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 120
Qy 963 AATGAAGAATAAACCAACACTGAC 986

```

Search completed: January 28, 2006, 02:37:37
Job time : 40.4546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 179.003 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-10
Perfect score: 2939
Sequence: 1 gcaacccgcagcctagcc.....gaataagctgaccatctaa 1661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/Cgn2_1/USPTO_epool_p/US09940235/runat_27012006.144218.27563/app_query.fasta_1.7708
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @Cgn_1_1_1359 @runat_27012006.144218.27563 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	67.5	440	1 STRP_STRSQ	P00779 streptococc
2	1955	66.5	414	2 Q53284_STRSQ	Q53284 streptococc
3	1943	66.1	436	2 Q6UK57_STRSQ	Q6UK57 streptococc
4	1942	66.1	440	1 STRP_STRS1	P10519 streptococc
5	1814	61.7	440	2 Q7X0Y1_STRPY	Q7X0Y1 streptococc
6	1802	61.3	440	2 Q7X0Y8_STRPY	Q7X0Y8 streptococc
7	1802	61.3	440	2 Q8K5R8_STRP3	Q8K5R8 streptococc
8	1794	61.0	440	1 STRP_STRPY	P10520 streptococc
9	1794	61.0	440	2 Q532X6_STRPY	Q532X6 streptococc
10	1787	60.8	440	2 Q7X0Y2_STRPY	Q7X0Y2 streptococc
11	1758	59.8	440	2 Q7X0Y7_STRPY	Q7X0Y7 streptococc
12	1758	59.8	440	2 Q5X9T6_STRP6	Q5X9T6 streptococc
13	1755	59.7	440	2 Q8NZA6_STRP8	Q8NZA6 streptococc
14	1723	58.6	440	2 Q7X0Y3_STRPY	Q7X0Y3 streptococc
15	1715	58.4	440	2 Q7X0Y0_STRPY	Q7X0Y0 streptococc
16	1715	58.4	440	2 Q7X0Y5_STRPY	Q7X0Y5 streptococc

17	1708	58.1	440	2	Q7X0X8_STRPY	Q7X0X8 streptococc
18	1686	57.4	440	2	Q7X0X9_STRPY	Q7X0X9 streptococc
19	1677	57.1	440	2	Q7X0X7_STRPY	Q7X0X7 streptococc
20	1670	56.8	440	2	Q57391_STRPY	Q57391 streptococc
21	1657	56.4	440	2	Q7X0Y6_STRPY	Q7X0Y6 streptococc
22	1652	56.2	440	1	STRQ_STRPY	P96471 streptococc
23	1624	55.3	432	2	Q7X0Y4_STRPY	Q7X0Y4 streptococc
24	728	24.8	141	2	Q7X0X2_STRSQ	Q7X0X2 streptococc
25	727	24.7	141	2	Q7X0X3_STRSQ	Q7X0X3 streptococc
26	724	24.6	141	2	Q7X0W1_STRSQ	Q7X0W1 streptococc
27	723	24.6	141	2	Q7WS87_STRSQ	Q7WS87 streptococc
28	717	24.4	141	2	Q7X0W3_STRSQ	Q7X0W3 streptococc
29	717	24.4	141	2	Q7X0X4_STRSQ	Q7X0X4 streptococc
30	715	24.3	141	2	Q7X0W5_STRSQ	Q7X0W5 streptococc
31	714	24.3	141	2	Q7X0X5_STRSQ	Q7X0X5 streptococc
32	712	24.2	141	2	Q7X0X6_STRSQ	Q7X0X6 streptococc
33	711	24.2	141	2	Q7X0W2_STRSQ	Q7X0W2 streptococc
34	711	24.2	141	2	Q7X0W9_STRSQ	Q7X0W9 streptococc
35	710	24.2	141	2	Q7X0W6_STRSQ	Q7X0W6 streptococc
36	707	24.1	141	2	Q7X0W0_STRSQ	Q7X0W0 streptococc
37	705	24.0	141	2	Q7X0V5_STRPY	Q7X0V5 streptococc
38	704	24.0	141	2	Q7X0X0_STRSQ	Q7X0X0 streptococc
39	701	23.9	141	2	Q7X0W7_STRSQ	Q7X0W7 streptococc
40	697	23.7	141	2	Q7X0W4_STRSQ	Q7X0W4 streptococc
41	696	23.7	141	2	Q7X0W8_STRSQ	Q7X0W8 streptococc
42	693	23.6	141	2	Q7X0X1_STRSQ	Q7X0X1 streptococc
43	690	23.5	141	2	Q7X0R2_STRPY	Q7X0R2 streptococc
44	686	23.3	141	2	Q7X0S4_STRPY	Q7X0S4 streptococc
45	684	23.3	141	2	Q7X0S0_STRPY	Q7X0S0 streptococc

ALIGNMENTS

RESULT 1

ID	STRP_STRSQ	STANDARD;	PRT;	440 AA.
AC	P00779;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Streptokinase C precursor.			
GN	Name=skc;			
OS	Streptococcus equisimilis.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=119602;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=H46A;			
RX	MEDLINE=85232082; PubMed=2989113; DOI=10.1016/0378-1119(85)90145-3;			
RA	Malke H., Roe B.A., Ferretti J.J.;			
RT	"Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A.";			
RL	Gene 34:357-362(1985).			
RN	[2]			
RP	PROTEIN SEQUENCE OF 27-440.			
RX	MEDLINE=83127125; PubMed=6760891;			
RA	Jackson K.W., Tang J.;			
RT	"Complete amino acid sequence of streptokinase and its homology with serine proteases.";			
RL	Biochemistry 21:6620-6625(1982).			
CC	-!- FUNCTION: This protein is not a protease, but it activates plasminogen by complexing with it. As a potential virulence factor, it is thought to prevent the formation of effective fibrin barriers around the site of infection, thereby contributing to the invasiveness of the cells.			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
CC	-----			

DR EMBL; K02986; AAA26974.1; -: Genomic DNA.
 DR EMBL; X72832; CAA51351.1; -: Genomic DNA.
 DR PIR; A09667; BZSO.
 DR PIR; A22801; A22801.
 DR PDB; 1BML; X-ray; C/D=38-399.
 DR PDB; 1L4D; X-ray; B=40-173.
 DR PDB; 1L4Z; X-ray; B=27-173.
 DR PDB; 1OOR; X-ray; A/B/C/D=177-314.
 DR SMR; P00779; 38-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW 3D-structure; Direct protein sequencing; Plasminogen activation;
 KW Signal; Virulence.
 FT SIGNAL 1 26
 FT CHAIN 27 440 Streptokinase C.
 FT VARIANT 195 195 L -> D.
 FT VARIANT 207 207 D -> L.
 FT CONFLICT 298 300 EKV -> LEVK (in Ref. 2).
 FT CONFLICT 438 438 N -> D (in Ref. 2).
 FT STRAND 180 180
 FT STRAND 184 194
 FT TURN 199 200
 FT TURN 203 204
 FT STRAND 205 205
 FT STRAND 209 214
 FT TURN 216 217
 FT STRAND 219 221
 FT HELIX 222 236
 FT TURN 238 239
 FT STRAND 240 252
 FT TURN 253 254
 FT STRAND 259 260
 FT STRAND 268 270
 FT STRAND 275 275
 FT STRAND 278 280
 FT TURN 282 284
 FT STRAND 287 289
 FT STRAND 292 304
 FT TURN 305 306
 SQ SEQUENCE 440 AA; 50140 MW; 8FCLF22648ACC77A CRC64;

Alignment Scores:
 Pred. No.: 2.58e-128 Length: 440
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.51% Indels: 0
 DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x STRP_STREQ (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCTGCTCAACAAACAGCCAAATGGTTGTT 242
 Db 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
 QY 243 AGCGTGTGCTGCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
 QY 303 GATCTAACATCAGACCTGCTCATGGAGGAAGAAAGACAGACAGCTTAAGTCCAAATCA 362
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACCATTTGCTGATGATGCGCGATGTCACATAAACTTGAGAAAGCTGACTTACTA 422
 Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 106
 QY 423 AAGGCTATTACAGAAACAATTGATCGCTACAGTCCACAGTAACGACGACTACTTTGAGGTC 482
 Db 107 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126

QY 483 ATTGATTTTCAAGCGATGCAACATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 542
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
 QY 543 AAAGATGGTTTCGGTAACCTTTGCCGACCCAACTGTCTCAAGAAATTTTGTAAAGCGGACAT 602
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
 QY 603 GTGGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGGAATCTGTTGATGTG 662
 Db 167 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
 QY 663 GAATATACTGTACAGTTTACTCCCTTAAACCTGTATGACGATTTTACACAGAGGCTCTCAA 722
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 QY 723 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTTACTA 782
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATAGCATTTATGAACGTCAC 842
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCACTCATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 902
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 903 TTTTACTTACCGTGTATAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286
 QY 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 1023 GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTGAACTGTTCACCATCAATAGCTT 1082
 Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1083 GATGTCGATACCAACGAATTTCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1142
 Db 327 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTATACGATCTCTCGTGATAAGCTAACTACTCTCAACAATCTC 1202
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1203 GATGCTTTTGGTATTATGAGTACTATACCTTAACCTGGAAGAGTAGAGGATATCAGATGAC 1262
 Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
 QY 1263 ACCAACCGTATACACCGTTTATATGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT 1322
 Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1323 CATTTAGCC 1331
 Db 407 HisLeuAla 409
 RESULT 2
 Q53284 STREQ
 ID Q53284 STREQ PRELIMINARY; PRT; 414 AA.
 AC Q53284;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE SKC-2.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]

QY	903	TTTACTTACCGTGTATAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	963
Db	241	PheThrTyrHisVallysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu	260
QY	963	AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG	1022
Db	261	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	280
QY	1023	GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTCACCATCAATACGTT	1082
Db	281	GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal	300
QY	1083	GATGTCATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTACGGAACGTAAC	1142
Db	301	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	320
QY	1143	TTAGACTTCAGAGATTTATACGATCTCTCGTGAAGGCTAAACTACTCTACACAATCTC	1202
Db	321	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	340
QY	1203	GATGCTTTTGGTATTATGGACTATACCTTAACTCGAAAAAGTAGAGGATATCAGATGAC	1262
Db	341	AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	360
QY	1263	ACCAACCGTATCATACCGTTTATATGGCAAGCGACCGCAAGGAGAGAGTGTAGCTAT	1322
Db	361	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr	380
QY	1323	CATTATGCC 1331	
Db	381	HisLeuAla 383	
RESULT 3			
Q6UK57_STREQ			
ID	Q6UK57_STREQ	PRELIMINARY;	PRT; 436 AA.
AC	Q6UK57;		
DT	05-JUL-2004	(T-EMBLrel. 27, Created)	
DT	05-JUL-2004	(T-EMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(T-EMBLrel. 27, Last annotation update)	
DE	Streptokinase.		
OS	Streptococcus equisimilis.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=119602;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=ATCC 35666;		
RA	Costa C.S., Torres F.A.G., Filho S.A.;		
RL	Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY168335; AAQ73571.1; -; Genomic_DNA.		
DR	SWR; Q6UK57; 59-394.		
DR	GO; GO:0016301; F:kinase activity; IEA.		
DR	DR GO; GO:0008243; F:plasmaogen activator activity; IEA.		
DR	GO; GO:0005515; F:protein binding; IEA.		
DR	InterPro; IPR004093; Staphylokinase.		
DR	InterPro; IPR008124; Streptokinase.		
DR	Pfam; PF02821; Staphylokinase; 3.		
DR	PRINTS; PR01753; STREPKINASE.		
KW	Kinase.		
SQ	SEQUENCE 436 AA; 49658 MW; CA2DD95F957D8F3C CRC64;		
Alignment Scores:			
Pred. No.:	1.73e-125	Length:	436
Score:	1943.00	Matches:	374
Percent Similarity:	98.43%	Conservative:	3
Best Local Similarity:	97.65%	Mismatches:	6
Query Match:	66.11%	Indels:	0
DB:	2	Gaps:	0
US-09-940-235-10 (1-1661) x Q6UK57_STREQ (1-436)			
QY	183	ATGTCGACCTGAGTGGCTGCTAGACCTTCATCTGTCACAAACAGCCAAATGCTGTT	242

Db 23 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 42
 QY AGCGTTGCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 43 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 62
 QY 303 GATCTAACATCACGACCTGCTCATGAGGAGAAACAGACAGCGCTTAAGTCCAAATCA 362
 Db 63 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 82
 QY 363 AAACATTGCTGATGATGCGCGGATGTCAATAAATTGAGAAAGCTGCTACTTA 422
 Db 83 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 102
 QY 423 AAGGCTATTCAAGACAAATTGCTGCTTAAGTCCAGACAGTAAACGACACTACTTTAGGTC 482
 Db 103 LysAlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 122
 QY 483 ATTGATTGTCAGCGATGCAACCATTTACTGATCGAAACGCAAGCTCTACTTTGCTGAC 542
 Db 123 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 142
 QY 543 AAAGATGTTGCGTAACTTCCGACCCCAACCTGTCCTCAAGAAATTTTGTAAAGCGGACAT 602
 Db 143 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 162
 QY 603 GTGGCGGTAGACCATATAAGAAACCAATACAAACCAACGCAAGTCTGTTGATGTG 662
 Db 163 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 182
 QY 663 CAATATATCTGACGTTTACTCCCTTAAACCTGATGACGATTTCAGACAGCTCTCAA 722
 Db 183 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 202
 QY 723 GATCTAAAGCTATTGAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTACTTA 782
 Db 203 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 222
 QY 783 GCTCAAGACAAAGCATTTTAAACAAACCCAGGCTTATGAGTATTAAGTATGACGTGAC 842
 Db 223 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 242
 QY 843 TCCTCAATCGTCACATCAGCAATGCAATTTCCGTACGATTTTACCAGTATGATCAAGAG 902
 Db 243 SerSerIleValThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGlu 262
 QY 903 TTTACTTACCGTGTAAAAATCGGAAACAAAGCTTATGAGTATCAATAAAAAATCTGCTGTG 962
 Db 263 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 282
 QY 963 AATGAGAAATAACAAACACTGATCTCTGAGAAATATATAGTCTCTTAAAAAGGG 1022
 Db 283 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 302
 QY 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTACCATCAATACGTT 1082
 Db 303 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 322
 QY 1083 GATGTCGATACCAACGAATTGCTAAAAAGTGAAGCTCTTAACAGCTAGCGCAACTAAC 1142
 Db 323 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 342
 QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAAGGCTTAAACTACTCTACCAACTCTC 1202
 Db 343 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 362
 QY 1203 GATGCTTTGTTGATTTATGACTATACCTTAACCTGAAAGTAGAGATATACAGTATGAC 1262
 Db 363 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 382
 QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1322

Db 383 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 402
 QY 1323 CATTTAGCC 1331
 Db 403 HisLeuAla 405
 RESULT 4
 STRP_STRS1
 ID STRP_STRS1 STANDARD; PRT; 440 AA.
 AC P10519;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 GN Streptokinase G precursor.
 GN Name=skg;
 OS Streptococcus sp. (strain 19909).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=69017;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RA MEDLINE=89160285; PubMed=2922269;
 RX Walter F., Siegel M., Malke H.;
 RT "Nucleotide sequence of the streptokinase gene from a group-G
 Streptococcus";
 RL Nucleic Acids Res. 17:1262-1262(1989).
 CC -!- FUNCTION: This protein is not a protease, but it activates
 plasminogen by complexing with it. As a potential virulence
 factor, it is thought to prevent the formation of effective fibrin
 barriers around the site of infection, thereby contributing to the
 invasiveness of the cells.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 CC EMBL; X13400; CAA31766.1; -; Genomic_DNA.
 DR PIR; S02723; S02723.
 DR HSSP; P00779; 1QOR.
 DR SMR; P10519; 63-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Plasminogen activation; Signal; Virulence.
 FT SIGNAL 1 26
 FT CHAIN 27 440 Streptokinase G.
 SQ SEQUENCE 440 AA; 50199 MW; 5521P8825F81B6EA CRC64;
 Alignment Scores:
 Pred. No.: 2,03e-125 Length: 440
 Score: 1942.00 Matches: 375
 Percent Similarity: 98.43% Conservative: 2
 Best Local Similarity: 97.91% Mismatches: 6
 Query Match: 66.08% Indels: 0
 DB: Gaps: 0
 US-09-940-235-10 (1-1661) x STRP_STRS1 (1-440)
 QY 183 ATTCGTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGCTGTT 242
 Db 27 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
 QY 243 AGCGTTGCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 302
 Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 66
 QY 303 GATCTAACATCACGACCTGCTCATGAGGAGAAACAGACAGCGCTTAAGTCCAAATCA 362
 Db 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86

QY	363	AAACCAATTTCGTACTGATAGTGGCCGGATGTGCATAAACCTTGAGAAAGCTGACTTACTA	422
Db	87	LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106
QY	423	AAGGGCTATTCAAGAACAATTGATCGCTAACGTCACAGTAAACGACGACGACTCTTTGAGGTC	482
Db	107	LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspSyrPheGluVal	126
QY	483	ATTGATTTTTGCCAAGCGATGCAACCATTACTGATCGAAAACGGCAAGGTCTACTTTGCTCAC	542
Db	127	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	146
QY	543	AAAGATGGTTCCGTTAACTTCCCGCAGCCCAACCTGTCCCAAGAAATTTTTCGTAAGCGGACAT	602
Db	147	LysAspGlySerValThrLeuProIleGlnProValGlnGluPheLeuLeuLysGlyHis	166
QY	603	GTCGCGGTTAGACCATATAAGAAAAACAATACACAAACCAACCAAGCGAAATCTGTCATGTG	662
Db	167	ValArgValArgProTyrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal	186
QY	663	GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCCAGACAGGTCTCAAA	722
Db	187	GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys	206
QY	723	GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA	782
Db	207	AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu	226
QY	783	GCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGCGCTATACGATTATGACGTGAC	842
Db	227	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp	246
QY	843	TCCTCAATCGTCACATCATGACAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG	902
Db	247	SerSerIleValThrHisAspAsnAspIlePheargThrIleLeuPrometAspGlnGlu	266
QY	903	TTTACTTACCGGTGTAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG	962
Db	267	PheThrTyrHisValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu	286
QY	963	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAAGGG	1022
Db	287	AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly	306
QY	1023	GAAGAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACGTTCACCATCAAATACGTT	1082
Db	307	GluLysProTyrAspProPheAspArgSerHisLysLysLeuPheThrIleLysTyrVal	326
QY	1083	GATGTTCGATACCAACGAATTCCTAAAAAGTCGACGAGCTCTTAAACAGCTAGCGAAGTAAAC	1142
Db	327	AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn	346
QY	1143	TTAGACTTCAGAGATTATACGATCTCTGATAGGCTAAAGCTAAACTCTCTACAACTCTC	1202
Db	347	LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu	366
QY	1203	GATGCTTTTGTATTATTCGACTATACCTTAACTCGAAAGTAGAGGATAATCACGATGAC	1262
Db	367	AspAlaPheGlyIleMetCaspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp	386
QY	1263	ACCAACCGTATCATACCGTTTATATGGCGAAGCGACCCGAAGGAGAGAAATGCTAGGTAT	1322
Db	387	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr	406
QY	1323	CATTAGCC	1331
Db	407	HisLeuAla	409
RESULT 5			
Q7X0Y1_STRPY PRELIMINARY; PRT; 440 AA.			
AC Q7X0Y1;			
DT 01-OCT-2003 (TrEMBLrel. 25, Created)			

```

DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13114;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D488;
RX PubMed=14579231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234137; AAP39957.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0V1; 63-398.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0008243; P:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4E6E9647043BAC CRC64;

Alignment Scores:
Pred. No.: 1.36e-116 Length: 440
Score: 1814.00 Matches: 347
Percent Similarity: 95.04% Conservative: 17
Best Local Similarity: 90.60% Mismatches: 19
Query Match: 61.72% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q7X0Y1 STREY (1-440)

```

183	QY	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT	242
	Db		
27	Db	IleAlaGlyTyrGlyTripLeuProAspArgProProlIleAsnAsnSerGlnLeuValVal	46
243	QY	AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC	302
	Db		
47	Db	SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPhePheGluIle	66
303	QY	GATCTTAACATCACGACCTGCTCATCGAGGAAAGACAGACGAGCGTTAAGTCCAAATCA	362
	Db		
67	Db	AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer	86
363	QY	AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATTGAGAAAGCTGACTTACTA	422
	Db		
87	Db	LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu	106
423	QY	AAGCTATTTCAGAACCAATTGATCGCTTAACTCCACAGTAACGACGACTCTTTGAGGTC	482
	Db		
107	Db	LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal	126
483	QY	ATTGATTTTGCACGCGATGCAACCAATTACTGATCGAAACGGCAAGCTTCTTTGCTGAC	542
	Db		
127	Db	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp	146
543	QY	AAAGATGGTTTCGGTAAACCTTCCCGACCCAAACCTGTCCAGAAATTTTGTGAAGCGGACAT	602
	Db		
147	Db	LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis	166
603	QY	GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAACCCAGCGAAATCTGTTGATGTG	662
	Db		
167	Db	ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal	186
663	QY	GAATATATCTGTACAGTTTACTTCCCTTAAACCTGTATGACGATTTACAGACCGAGTCTCAA	722

RESULT 5
Q7X0Y1 S
ID Q7X
AC Q7X
DT 01-

Db 187 LysThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 723 GATACCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 782
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 783 GCTCAAGCAACAAGCAATTTAAACAAACCCAGGCTATACGATTTATGATCAACGTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyThrIleThrLysGluArgAsp 246
QY 843 TCCTCAATCGTCACTCATGACATGACATTTCCGTACGATTTTACCAATGATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAATAATCGGAACCAAGCTTATAGGATCAATAAAATACTGTGCTG 962
Db 267 PheThrTyrgValLysAspArgGluGlnAlaTyrgLysLysSerGlyLeu 286
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAGGG 1022
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrrIleLeuLysLysGly 306
QY 1023 GAAAGCCGTATGATCCCTTTGATCCAGTCACTTGAACCTGTTCAACCAATACGTT 1082
Db 307 GluSerProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrVal 326
QY 1083 GATGTCGATACCAAGAAATGCTTAAAGATGAGCAGCTCTTAACAGCTAGCGAACCTAAC 1142
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGCTGATGAGGCTAAACTACTCTACCAACAATCTC 1202
Db 347 LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu 366
QY 1203 GATGCTTTGGTATTATGACATATACCTTAACCTGAAAAAGTAGAGGATATACGATGAC 1262
Db 367 AspAlaPheAspIleMetAspTyrrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1263 ACCACCGTATCATACCGTATTATAGGCAAGCGACCGACCGAGGAGATGCTACTAT 1322
Db 387 AsnAsnArgIleValThrValTyrrMetGlyLysArgProLysGlyAlaLysGlySerTyrr 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 6

Q7X0Y8 STRPY
ID Q7X0Y8 STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y8
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=88-019;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL: AY234129; RAP39949.1; -; Genomic_DNA.
DR HSP; Q53284; IC4P.
DR SMR; Q7X0Y8; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.

DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SSSEQUENCE 440 AA; 49885 MW; 69DB44F4026E3975 CRC64;

Alignment Scores:

Pred. No.: 9, 12e-116 Length: 440
Score: 1802.00 Matches: 347
Percent Similarity: 94.52% Conservative: 15
Best Local Similarity: 90.60% Mismatches: 21
Query Match: 61.31% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q7X0Y8_STRPY (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCCAATTTGTTT 242
Db 27 IleAlaGlyTyrrGlyTyrrLeuProAspArgProValAsnAsnSerGlnLeuVal 46
QY 243 AGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 302
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAGCTTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnHisAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAACTTGAGAAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeu 106
QY 423 AAGGCTATTCAAGAACAAATTGATCGCTTAACGTCCACAGTAAACGACGACTCTTTGAGTC 482
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrrPheGluVal 126
QY 483 ATTGATTTTCAAGCGATGCAACCATTTACTGTGTCGAAACCGCAAGGTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrrPheAlaAsp 146
QY 543 AAAGATGTTTCCGTAACTTTCGCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 602
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 603 GTCGCGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
Db 167 ValArgValArgProTyrrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 663 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 722
Db 187 LysTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 723 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
Db 207 AspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 783 GCTCAAGCAACAAGCAATTTTAAACAAACCCAGGCTTATAGGATCAATAAAATACTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyThrIleTyrrGluArgAsp 246
QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAAAAATCGGAACCAAGCTTATAGGATCAATAAAATACTGCTGCTG 962
Db 267 PheThrTyrgValLysAspArgGluGlnAlaTyrgLysLysSerGlyLeu 286
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAGGG 1022
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrrIleLeuLysLysGly 306

Db 327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTATACGATCTCGTGAAGCTAAAGCTACTCTCAACAATCTC 1202
 |||||
 Db 347 LeuAspPheArgAspLeuThrAspProCysAspLysAlaLysLeuLeuThrAsnAsnLeu 366
 |||||
 QY 1203 GATGCTTTGGTATTATGACTATACCTTTAACTGGAAAAAGTAGAGGATAATCAATCGATGAC 1262
 |||||
 Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
 |||||
 QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCAGCCGAGGAGAGAAATGCTAGCTAT 1322
 |||||
 Db 387 AsnAsnArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
 |||||
 QY 1323 CATTTAGCC 1331
 |||||
 Db 407 HisLeuAla 409

RESULT 8

STRP_STRPY
 ID _STRP_STRPY STANDARD; PRT; 440 AA.
 AC P10520;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Streptokinase A precursor.
 GN Name=ska; OrderedLocusNames=SPY1979;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SF130/13 / Serotype M1;
 RX MEDLINE=89160264; PubMed=2646590;
 RA Walter F., Siegel M., Malke H.;
 RT "Nucleotide sequence of the streptokinase gene from a Streptococcus
 pyogenes type 1 strain."
 RL Nucleic Acids Res. 17:1261-1261(1989).
 RN [2]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 CC -1- FUNCTION: This protein is not a protease, but it activates
 plasminogen by complexing with it. As a potential virulence
 factor, it is thought to prevent the formation of effective fibrin
 barriers around the site of infection, thereby contributing to the
 invasiveness of the cells.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

 EMBL; X13399; CA311765.1; -; Genomic DNA.
 EMBL; AE006620; AAK34665.1; -; Genomic DNA.
 DR PIR; S02724; S02724.
 DR HSSP; Q53284; 1C4P.
 DR SRR; P10520; 63-398.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 DR Complete proteome; plasminogen activation; Signal; Virulence.
 FT SIGNAL 1 26
 CHAIN 27 440 Streptokinase A.

FT CONFLICT 163 163 L -> V (in Ref. 1).
 FT CONFLICT 345 345 R -> G (in Ref. 1).
 FT CONFLICT 373 373 D -> N (in Ref. 1).
 FT CONFLICT 428 428 D -> Y (in Ref. 1).
 FT CONFLICT 438 438 K -> N (in Ref. 1).
 SQ SEQUENCE 440 AA; 49924 MW; D6227BF040B758DB CRC64;
 Alignment Scores:
 Pred. No.: 3 25e-115 Length: 440
 Score: 1794.00 Matches: 341
 Percent Similarity: 93.99% Conservative: 19
 Best Local Similarity: 89.03% Mismatches: 23
 Query Match: 61.04% Indels: 0
 DB: 1 Gaps: 0

US-09-940-235-10 (1-1661) x STRP_STRPY (1-440)

QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCCAATGTTGTT 242
 |||||
 Db 27 IleAlaGlyTyrGlyTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal 46
 |||||
 QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 302
 |||||
 Db 47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPhePheGluIle 66
 |||||
 QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACGAAGCTTTAAGTCCAAAATCA 362
 |||||
 Db 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
 |||||
 QY 363 AAACCAATTTGCTACTGATGTCGGCGATGTCATATAAACTTGAGAAAGCTGACTTACTA 422
 |||||
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
 |||||
 QY 423 AAGGCTATTCAAGAACAAATTGATCGCTTAACGTCACAGTAAACGACGACTTTGAGGTC 482
 |||||
 Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
 |||||
 QY 483 ATTGATTTTCAGACGATGCAACATTACTGTCGAACGGCAAGCTTACTTTGCTGAC 542
 |||||
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspAsnGlyLysValTyrPheAlaAsp 146
 |||||
 QY 543 AAAGATGTTTCGGTAACTTCGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGGACAT 602
 |||||
 Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuLysGlyHis 166
 |||||
 QY 603 GTCGCGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 662
 |||||
 Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
 |||||
 QY 663 GAATATACGTACAGTTTACTCCCTTAAACCCGTGATGACGATTCAGACGAGGTCTCAA 722
 |||||
 Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
 |||||
 QY 723 GATACTAAGCTATTGAAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 782
 |||||
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
 |||||
 QY 783 GCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTCGAC 842
 |||||
 Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
 |||||
 QY 843 TCCTCAATCGTCACTCATGACATGATTTTCGTCGATAGATTTTACCAATGGATCAAGAG 902
 |||||
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 |||||
 QY 903 TTTTACTTTACCGTGTAAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
 |||||
 Db 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluLeuAsnProLysThrGlyLe 286
 |||||
 QY 963 AATGAAGAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 |||||
 Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306

```

QY 1023 GAAAGCCGTATGATCCCTTTGATCGAGTCACCTTGAAACTGTTCAACATCAATACGTT 1082
Db |||||
QY 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Db |||||
QY 1083 GATGTCGATACCAACCAATTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAAGTAAC 1142
Db |||||
QY 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Db |||||
QY 1143 TTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAACTCTCTACAACAATCTC 1202
Db |||||
QY 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Db |||||
QY 1203 GATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1262
Db |||||
QY 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Db |||||
QY 1263 ACCAACCGTATCATACCTTTATATGCGCAAGCGAGCCGCAAGAGAGAGATGCTAGCTAT 1322
Db |||||
QY 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
Db |||||
QY 1323 CATTTAGCC 1331
Db |||||
QY 407 HisLeuAla 409
Db |||||
RESULT 9
ID Q53ZX6_STRPY PRELIMINARY; PRT; 440 AA.
AC Q53ZX6.
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86-779.
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234128; AAP39948.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;
Alignment Scores:
Pred. No.: 3.25e-115 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.99% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 61.04% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-10 (1-1661) x Q53ZX6_STRPY (1-440)
QY 183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACCAACAGCCAATTTGGTTGTT 242
Db |||||
QY 27 IleAlaGlyTyrGlyTyrLeuProAspArgProPheIleAsnAsnSerGlnLeuValVal 46
Db |||||
QY 243 AGCGTGTCTGGTCTGTTGGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db |||||
QY 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
Db |||||
QY 303 GATCTAAATCAACGACCTGCTCATGGAGGAAGAGACAGACGAGCTTAAGTCCAAATCA 362
Db |||||
QY 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
Db |||||
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTGAGAAAGCTGACTTACTA 422
Db |||||

```

```

Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAACTTCATCGCTAACTGTCACACTAGCAGCAGCTACTTTGAGTGC 482
Db |||||
QY 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Db |||||
QY 483 ATTGATTTTGCACGATGCAACCATCTGATCGAAACGGCAGGCTCTACTTTGCTGAC 542
Db |||||
QY 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
Db |||||
QY 543 AAAATGTTTCGGTAACCTTCCCGACCCCACTGCTCAAGAAATTTTCTAAGCGGACAT 602
Db |||||
QY 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLysGlyHis 166
Db |||||
QY 603 GTGCGCTAGACCATATAAGAAACCAATACAAACCAACGCGAAATCTGTCATGTCG 662
Db |||||
QY 167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
Db |||||
QY 663 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 722
Db |||||
QY 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
Db |||||
QY 723 GATTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 782
Db |||||
QY 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
Db |||||
QY 783 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTATCAACGTCAC 842
Db |||||
QY 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
Db |||||
QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTGACATTTTACCAATGGATCAAGAG 902
Db |||||
QY 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetCaspGlnGlu 266
Db |||||
QY 903 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
Db |||||
QY 267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
Db |||||
QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATAGTCTTAAAAAAGG 1022
Db |||||
QY 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrTyrValLeuLysGlnGly 306
Db |||||
QY 1023 GAAAGCCGTATGATCCCTTTGATCGAGTCACCTTGAAACTGTTCAACATCAATACGTT 1082
Db |||||
QY 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Db |||||
QY 1083 GATGTCGATACCAACCAATTGCTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAC 1142
Db |||||
QY 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Db |||||
QY 1143 TTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAACTCTCTACAACAATCTC 1202
Db |||||
QY 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Db |||||
QY 1203 GATGCTTTTGGTATTATGAGCTATACCTTAACCTGGAAGAGTAGAGGATAATCAGATGAC 1262
Db |||||
QY 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Db |||||
QY 1263 ACCAACCGTATCATACCTTTATATGCGCAAGCGAGCCGCAAGAGAGAGATGCTAGCTAT 1322
Db |||||
QY 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
Db |||||
QY 1323 CATTTAGCC 1331
Db |||||
QY 407 HisLeuAla 409
Db |||||
RESULT 10
QYQY2_STRPY
ID Q7X0Y2_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y2.
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

```

DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D306;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234136; AAP39956.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y7; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50188 MW; 4CB42AC2A9062C2E CRC64;

Alignment Scores:
Pred. No.: 9,87e-115 Length: 440
Score: 1787.00 Matches: 342
Percent Similarity: 94.26% Conservative: 19
Best Local Similarity: 89.30% Mismatches: 22
Query Match: 60.80% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q7X0Y2_STRPY (1-440)

QY 183 ATGCTGACCTGAGTGGCTGCTAGACCGTCTGTCACGACCAACCAAGCAATTTGGTTGTT 242
Db 27 IleAlaGlyTyrGlyTyrPleuProAspArgProFoilEasAsnSerGlnLeuValVal 46

QY 243 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGCAATAGTCTTAAATTTTTTGAATC 302
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66

QY 303 GATCTAACATCACGACCTGCTCATGGAGGAAGACAGACCAAGCTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86

QY 363 AAACCATTTGCTACTGATGAGTGGCGGATGTACATAACTTGAGAAAGCTGACTTACTA 422
Db 87 LysProPheAlaThrAspAsnSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106

QY 423 AAGCGTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGTTC 482
Db 107 LysAlaIleGlnLysGlnLeuIleAlaValHisSerAsnAspGlyTyrPheGluVal 126

QY 483 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGCTTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 543 AAAGATGCTTGGTTAACTTGGCGCCGCACTGCTCCAGAAATTTTGTACGCGACAT 502
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuThrGlyHis 166

QY 603 GTGCGCGTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCAAACTGTTGATGTG 662
Db 167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186

QY 663 GAATATAGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAAA 722
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206

QY 723 GATACTAAGCTATTGAAAAACACATAGCTATCGTGACACCATCATCATCTCAAGAAATTACTA 782
Db 207 AspThrLysLeuLeuLysLysLeuAlaIleGlyAspThrValThrSerGlnGluLeuLeu 226

QY 783 GCTCAAGCACAAGAGCTTTTAAACAAACACCCAGCGCTATACGATTTATGAAACGTTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 843 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 247 SerSerIleValThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 903 TTTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 962
Db 267 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 286

QY 963 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysArg 306

QY 1023 GAAAGCCGTATCATCCCTTTGATCGCAGTCACCTTGAACCTGTTCAACCATCAATAGTT 1082
Db 307 GluLysProTyrAspPropheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326

QY 1083 GATGTCGATACCAACCAATTTGCTAAAAAGTGACAGCTCTTAAACAGCTAGCAGCACTAAC 1142
Db 327 AspValAsnThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346

QY 1143 TTAGACTTCAGAGATTTATACGATCTCTCGTGATAGGCTAAAGCTAACTCTTACAACATCTC 1202
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366

QY 1203 GATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAGGATATACGATGAC 1262
Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386

QY 1263 ACCAACCGTATCATACCAACCGTTTATATGGCAACGACCGACCGAGAGAGATGCTAGCTAT 1322
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406

QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 11
Q7X0Y7_STRPY
ID Q7X0Y7_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IRP112;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234130; AAP39950.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y7; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.

DR Pfam: PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW KINASE.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 9,86e-113 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservativeness: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 59.82% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x QYX0Y7_STRPY (1-440)

QY 183 ATTGCTGGAGCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCAATTTGTTGTT 242
DB 27 IleAlaGlyTyrGlyTrpLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 243 AGCGTTGCTGGTACTGTGTAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 302
DB 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66

QY 303 GATCTAACATCAGCACTGCTCATGGAGGAGAAACAGACAGCAAGCTTAAGTCCAAATCA 362
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86

QY 363 AAACCATTTGCTACTGATGTGGCGATCTCATATAAATCTTGAGAAGCTGACTTACTA 422
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeu 106

QY 423 AAGGCTATTCAAGAACAAATGTATGCTGCTAAGTCCACAGTACACGCACTACTTTGAGTC 482
DB 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126

QY 483 ATTGATTTTGACAGCAATCAACCATTTACTGATCGGAAACCGCAAGTCTACTTTGTCGAC 542
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146

QY 543 AAAGATGGTTCGTACCTTGGCCGACCACTGCTCCAGAAATTTTGTCTAAGCGGCAT 602
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166

QY 603 GTGCGGTGTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGTATGTG 662
DB 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186

QY 663 GAATATACGTACAGTTTACTCCCTTAAACCTGTATGACGATTCAGACAGGTCTCTCAA 722
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206

QY 723 GATACCTAAGCTATTGAAACACTAGTATCGGTGACACCATCATCTCAAGAAATCTACTA 782
DB 207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226

QY 783 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGACCTGAC 842
DB 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyrThrIleTyrGluArgAsp 246

QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266

QY 903 TTTACTACCGTGTAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGCTGCTG 962
DB 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286

QY 963 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGCTCTTAAAAAGGG 1022
DB 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrThrValLeuLysLysGly 306

QY 1023 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTCAAACTGTTTACCACATCAATACGTT 1082
DB 1023 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTCAAACTGTTTACCACATCAATACGTT 1082

DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326

QY 1083 GATGTCGATACCAACGAATTCGTAATAAGTCAGCAGCTCTTAACAGCTAGCGAACGTAAAC 1142
DB 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346

QY 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTACTCTACAAACAATCTC 1202
DB 347 LeuAspPheAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366

QY 1203 GATGCTTTGTTATATGACTATACCTTAACCTGGAAGAGTAGAGATAATCAGCATGAC 1262
DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386

QY 1263 ACCAACGCTATCATACCGTTTATATGCGCAAGCGACCGGAGGAGAGAGAGTCTAGCTAT 1322
DB 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406

QY 1323 CATTTAGCC 1331
DB 407 HisLeuAla 409

RESULT 12
QX9T6_STRP6
ID QX9T6_STRP6 PRELIMINARY; PRT; 440 AA.
AC QX9T6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Streptokinase (SC 3.4.-.-).
GN OrderedLocustNames=M6_Spy1692;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS10394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.E.,
RA Voyich J.M., Deleo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
metagenome: complete genome sequence of a macrolide-resistant serotype
M6 strain."
RL J. Infect. Dis. 190:727-738 (2004).
DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
DR SMR; QX9T6; 63-398.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Hydrolase; Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;

Alignment Scores:

Pred. No.: 9,86e-113 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservativeness: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 59.82% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x QX9T6_STRP6 (1-440)

QY 183 ATTGCTGGAGCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGCAATTTGTTGTT 242
DB 27 IleAlaGlyTyrGlyTrpLeuLeuAspArgProValAsnAsnSerGlnLeuVal 46

QY 243 AGCGTTGCTGGTACTGTTGAGGGGACGATCAAGACATTAGTCTTAAATTTTGAATC 302


```

Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
QY 303 GATCTAAACATACGACCTCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTGCATATAAACTTCAGAAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 482
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 483 ATTGATTGTCGACGATGACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 543 AAAGATGGTTCCGTAACCTTCGCGACCCCACTCTCCAAAGATTTTTCGCTAAGCGGACAT 602
Db 147 LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuArgGlyHis 166
QY 603 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAACGCGAAATCTGTGTGATGTC 662
Db 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186
QY 663 GAATATACGTACAGTTTACTCCCTTAACCTTCATGATGACGATTTGACAGGCTCTCAAA 722
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206
QY 723 GATCTACGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATACTA 782
Db 207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226
QY 783 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246
QY 843 TCCTCAATCGTCATCATGACAAATGACATTTCCGTCACGATTTTACCAATGAAATCAAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 903 TTTACTTACCGTGTAAATAACCGGAACAGCTTATAGGATCAATAAATAATCGCTCTG 962
Db 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286
QY 963 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1022
Db 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
QY 1023 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAACACTGTTCACCAATCAATACGTT 1082
Db 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
QY 1083 GATGTCGATACCAACGAATGCTAAAGTAGGAGCAGCTCTTAAACAGCTAGCGAACCTAAC 1142
Db 327 AspValAsnThrAsnLysLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTTCAGAGATTATACGATCTCTGATGATAGGCTTAACTACTCTACCAATCTTC 1202
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1203 GATGCTTTGTTATGATGACTATACCTTAACTGGAAGAGTAGAGGATATCACGATGAC 1262
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGCCGAGGAGAGAGATGCTAGCTAT 1322
Db 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 406
QY 1323 CATTTAGCC 1331

```

```

Db 407 HisLeuAla 409
RESULT 13
QBNZA6_STRP8
ID QBNZA6_STRP8 PRELIMINARY; PRT; 440 AA.
AC QBNZA6;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Streptokinase.
GN OrderedLocustNames=spyM18_2042;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
NCBI_TaxID=301451;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F., Zhang Q.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Musser J.M.;
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AL010107; ALU98517.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; QBNZA6; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Streptokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Kinase.
SQ SEQUENCE 440 AA; 50382 MW; 47DAB35577EBB88E CRC64;

Alignment Scores:
Pred. No.: 1,59e-112 Length: 440
Score: 1755.00 Matches: 334
Percent Similarity: 93.73% Conservativeness: 25
Best Local Similarity: 87.21% Mismatches: 24
Query Match: 59.71% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x QBNZA6_STRP8 (1-440)
QY 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGTTGTT 242
Db 27 IleAlaGlyTyrGluTyrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 243 AGCGTGTCTGCTACTCTGAGGGGACGATCAAGACATAGTCTTAAATTTTGAATC 302
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPhePheGluIle 66
QY 303 GATCTAAACATACGACCTCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTGCTACTGATAGTGGCGGATGTGCATATAAACTTCAGAAAGCTGACTACTA 422
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY 423 AAGGCTATTCAAGAACAAATGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 482
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 483 ATTGATTGTCGACGATGACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnIleTyrPheAlaAsn 146

```


QY 543 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
 Db ::
 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
 QY 603 GTGCGGTAGACCATATAAGAAACCAACCAATACAAACCAACGGAATCTGTTGATGTG 662
 Db ::
 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
 QY 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATCAGCATTTTCAGACAGGTCCTCAA 722
 Db ::
 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
 QY 723 GATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
 Db ::
 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db ::
 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisSerAspTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902
 Db ::
 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 903 TTTACTTACCGTGTAAAAATCGGGAACCAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
 Db ::
 267 PheThrTyrHisIleLysAspArgGluGlnAlaTyrGlyIleAsnLysSerGlyGln 286
 QY 963 ANTGAAGAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
 Db ::
 287 GluGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 306
 QY 1023 GAAAGCCCTATGATCCCTTGTGATCGCAGTCACTTGAACCTGTTCAACATCAATAGCTT 1082
 Db ::
 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
 QY 1083 GATGTCGATACCAACGAATTCCTTAAAGAGTACGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
 Db ::
 327 AspValAsnThrAsnLysLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
 QY 1143 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGCTTAACTACTCTTACAACTCTC 1202
 Db ::
 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1203 GATGCTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATATACGATGAC 1262
 Db ::
 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
 QY 1263 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCGAGGAGAGATGCTAGCTAT 1322
 Db ::
 387 AsnAsnArgValValThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 406
 QY 1323 CATTTAGCC 1331
 Db ::::::::::::::::::
 407 HisLeuAla 409

RESULT 14
 Q7X0Y3_STRPY PRELIMINARY; PRT; 440 AA.
 ID Q7X0Y3_STRPY PRELIMINARY;
 AC Q7X0Y3;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ALAB49;

RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Bessen D.E.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL; AY234134; AAPJ9954.1; -; Genomic_DNA.
 DR HSSP; Q53284; 1C4P.
 DR DR SMR; Q7X0Y3; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasmaingogen activator activity; IEA.
 DR GO; GO:0005545; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50342 MW; B3C60BAA50DB2C39 CRC64;
 Alignment Scores:
 Pred. No.: 2.55e-110 Length: 440
 Score: 1723.00 Matches: 327
 Percent Similarity: 92.43% Conservatives: 27
 Best Local Similarity: 85.38% Mismatches: 29
 Query Match: 58.63% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-10 (1-1661) x Q7X0Y3_STRPY (1-440)
 QY 183 ATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTTGTTGTT 242
 Db ::
 27 IleAlaGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
 QY 243 AGCGTCTGCTACTGTTGAGGACGAATCAAGACATTAAGTCTTAATTTTGTGAATC 302
 Db ::
 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPheGluIle 66
 QY 303 GATCTAATCATCACACCTGCTCATGGAGGAAAGACAGCAAGCTTAAGTCCAAATCA 362
 Db ::
 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
 QY 363 AAACATTTGCTACTGATGATGGCGCATGTGCATATAACTTGAGAAAGCTGACTTACTA 422
 Db ::
 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeu 106
 QY 423 AAGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAAACGACGACTCTTTGAGTGC 482
 Db ::
 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
 QY 483 ATTGATTTTCACAGCATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 542
 Db ::
 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
 QY 543 AAAGATGGTTCGGTAACCTTGGCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
 Db ::
 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgGlyHis 166
 QY 603 GTGCGCTTAGACCATATAAGAAAAACCAATACAAACCAACGGAATCTGTTGATGTG 662
 Db ::
 167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
 QY 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATCAGCATTTTCAGACAGGTCCTCAA 722
 Db ::
 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
 QY 723 GATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 782
 Db ::
 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 783 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
 Db ::
 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246
 QY 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 902

```

|||||
247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
QY 903 TTTACTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGCTCG 962
Db 267 PheThrTyHisValLysAsnArgGluGlnAlaTyGlnAsnAspAsnLysThrGlyLeu 286
QY 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTTACCTCTTAAAAAGGG 1022
Db 287 LysLysGluThrLysAsnThrAspLeuIleSerGluLysTyTyIleLeuLysLysGly 306
QY 1023 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACACTGTTCACCAATCAATACGTT 1082
Db 307 GluLysProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
QY 1083 GATGTCGATACCAACGAATTGCTAAAGTAGAGACTCTTAAACAGCTAGCGAACGTAAC 1142
Db 327 AspValAspThrLysAspLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAGGCTAAGCTTAACACTCTTACAACTCTC 1202
Db 347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 366
QY 1203 GATGCTTTTGGTATTATGAGCTATACCTTAACTCGAAAGTAGAGGATAATCACGATGAC 1262
Db 367 AspAlaPheAspIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY 1263 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAATGCTAGTAT 1322
Db 387 AsnAsnArgValValThrValTyMetGlyLysArgProLysGlyAlaLysGlySerTy 406
QY 1323 CATTTAGCC 1331
Db 407 HisLeuAla 409

RESULT 15
Q7X0Y0 STRPY
ID Q7X0Y0_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y0;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=d633;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234139; AAP39958.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; O7X0Y0; 63-398.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0008243; P:plasminogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 50143 MW; 4ACB2C29349D680C CRC64;

Alignment Scores:
Pred. No.: 9 09e-110 Length: 440
Score: 1715.00 Matches: 325
Percent Similarity: 91.91% Conservative: 27

```

```

Best Local Similarity: 84.86% Mismatches: 31
Query Match: 58.35% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-10 (1-1661) x Q7X0Y0_STRPY (1-440)
QY 183 ATTGCTGGAGCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTTGTT 242
Db 27 IleAlaGlyProGluTrpLeuLeuGlyArgProProValAsnAsnSerGlnLeuValVal 46
QY 243 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 302
Db 47 SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
QY 303 GATCTAACATCAGCACTGCTCATGAGGAGAAACAGACAGCAAGCTTTAAGTCCAAATCA 362
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 86
QY 363 AAACCATTTGCTACTGATGCGCGATGCTACATAAACTTCGAGAAAGCTGACTTACTA 422
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeu 106
QY 423 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACCACTACTTTGAGGTC 482
Db 107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyThrPheGluVal 126
QY 483 ATTGATTTTCAAGCGATGCAACATTACTGTATCGAAACGGCAAGGTCCTACTTTGCTGAC 542
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAspAsnIleTyThrPheAlaAsn 146
QY 543 AAAGATGGTTCGGTAAACCTTGGCCGACCACTGTCACAGAATTTTGTAAAGCGGACAT 602
Db 147 GlnAspGlySerValThrLeuProThrGlnProIleGlnPheLeuLeuArgIleHis 166
QY 603 GTGCGGCTTAGACCATATAAAGAAAAACCAATAACAAACCAACGAAATCTGTGATGTG 662
Db 167 ValArgValArgProTyLysGluLysProIleGlnThrProAlaLysSerValAspIle 186
QY 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGTAGACGATTTACAGACAGGTCCTCAA 722
Db 187 ArgTyThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys 206
QY 723 GATACTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCATCTCAAGATTA 782
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAsnThrIleThrSerGlnGluLeu 226
QY 783 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACAGTTTATGAACGTGAC 842
Db 227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyThrIleTyGluArgAsp 246
QY 843 TCCTCAATCTCTCATCATCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 902
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
QY 903 TTTACTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGCTCG 962
Db 267 PheThrTyHisValLysAsnArgGluGlnAlaTyLysLysLeuLysLeuPheThrIleLysTyVal 286
QY 963 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 1022
Db 287 LysGluLysThrAsnAsnThrAspLeuIleSerGluLysTyTyIleLeuLysLysGly 306
QY 1023 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACACTGTTCACCAATCAATACGTT 1082
Db 307 GluGluProTyAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 326
QY 1083 GATGTCGATACCAACGAATTGCTAAAGTAGAGAGCTCTTAAACAGCTAGCGAACGTAAC 1142
Db 327 AspValAspThrLysAlaLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1143 TTAGACTTCAGAGATTATACGATCCTCGTGATAGGCTAAGCTTAACACTCTTACAACTCTC 1202
Db 347 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 366

```

QY	1203	GATGCTTTTGGTATTATGGA	TATACCTTA	CTGAAAGTAGAGGATAATC	ACGATGAC	1262
Db	367	AspAlaPheAspIleMetAsp	TyrThrLeu	ThrGlyLysValGluasp	AsnHisAspLys	386
QY	1263	ACCAACCGTATCATACCGTT	TATATGCGCAAGCGACCCG	AGAGAGAAATGCTAGCTAT		1322
Db	387	AsnAsnArgValValThrVal	TyrMetGlyLysArgProl	ysGlyAlaLysGlySer	Tyr	406
QY	1323	CATTTAGCC				1331
Db	407	HisLeuAla				409

Search completed: January 28, 2006, 02:33:51
 Job time : 195.003 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:53:18 ; Search time 9059.25 Seconds
(without alignments)
11181.386 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tcgcttcacgttcgtcgcg.....ataacctaacgacaaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues 11766282
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_ses:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245.8	69.9	2385	6	AR143998 Sequence
2	1244.4	69.8	2568	1	STRSKC
3	1244.4	69.8	2568	6	A04926
4	1244.4	69.8	8931	1	SEDEXB
5	1240.4	69.6	1242	6	AR144000
6	1239.6	69.6	2568	6	E00522
7	1237	69.4	1401	6	I05204
8	1230.4	69.0	2566	6	AR068768
9	1223.4	68.7	1458	6	A20027
10	1223.4	68.7	1458	6	I13215
11	1222.6	68.6	1245	1	S46536
12	1222.6	68.6	7057	6	CQ797820
13	1220.4	68.5	1335	6	A20006
14	1220.4	68.5	1335	6	I13194
15	1220.4	68.5	1512	6	A20016
16	1220.4	68.5	1512	6	I13204
17	1220.2	68.5	2589	6	I13209
18	1219.4	68.4	1257	6	A20015

19	1219.4	68.4	1257	6	I13203	I13203 Sequence 25
20	1219.4	68.4	1317	6	A20009	A20009 SEQ ID NO:
21	1219.4	68.4	1317	6	I13197	I13197 Sequence 18
22	1216.4	68.3	1467	6	A20030	A20030 SEQ ID NO:
23	1216.4	68.3	1467	6	I13218	I13218 Sequence 46
24	1216.4	68.3	2588	6	A20021	A20021 SEQ ID NO:
25	1212.4	68.0	1473	1	SGSKG	XI3400 Streptococc
26	1211.4	68.0	1407	6	E01413	E01413 DNA sequenc
27	1210.8	67.9	1311	1	AY368335	AY368335 Streptoco
28	1184	66.4	1209	6	AR175891	AR175891 Sequence
29	1184	66.4	1209	6	AX030315	AX030315 Sequence
30	1181.4	66.3	1245	6	AR175892	AR175892 Sequence
31	1181.4	66.3	1245	6	AX030316	AX030316 Sequence
32	1095.6	61.5	50354	1	AE014169	AE014169 Streptoco
33	1095.6	61.5	110000	1	BA000034_16	Continuation (17 o
34	1095.6	61.5	110000	1	BA000034_17	Continuation (18 o
35	1094	61.4	1323	1	AY234129	AY234129 Streptoco
36	1094	61.4	1323	1	AY234130	AY234130 Streptoco
37	1094	61.4	1323	1	AY234137	AY234137 Streptoco
38	1094	61.4	110000	1	CP000003_16	Continuation (17 o
39	1093.6	61.4	1122	6	AR175893	AR175893 Sequence
40	1093.6	61.4	1122	6	AX030317	AX030317 Sequence
41	1093.6	61.4	1158	6	AR175894	AR175894 Sequence
42	1093.6	61.4	1158	6	AX030318	AX030318 Sequence
43	1089.2	61.1	1323	1	AY234136	AY234136 Streptoco
44	1081.2	60.7	1323	1	AY234128	AY234128 Streptoco
45	1081.2	60.7	1323	6	CQ872087	CQ872087 Sequence

ALIGNMENTS

RESULT 1
LOCUS AR143998 2385 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6210667.
ACCESSION AR143998
VERSION AR143998.1 GI:15105865
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2385)
AUTHORS Reed G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 1 03-APR-2001;
FEATURES Location/Qualifiers
source
1..2385
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	69.9%;	Score 1245.8;	DB 6;	Length 2385;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1250;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	523	TTACCCGATGTTCTATTGCTGGACCTGAGTGGTGTAGACCTGTCATCTGTCAACAAC	582	
Db	1129	TCCATCGAGGTAGGATTGCTGGACCTGAGTGGTGTAGACCTGTCATCTGTCAACAAC	1188	
Qy	583	AGCCAATTGTTGTTAGCTGCTGTTGAGGGGACGATCAAGACATTAGTCTT	642	
Db	1189	AGCCAATTGTTGTTAGCTGCTGTTGAGGGGACGATCAAGACATTAGTCTT	1248	
Qy	643	AAATTTTGAATTCGATCTAAACATCAGACCTCTCATGAGGAAAGACAGACCAAGGC	702	
Db	1249	AAATTTTGAATTCGATCTAAACATCAGACCTCTCATGAGGAAAGACAGACCAAGGC	1308	
Qy	703	TTAAGTCCAAAATCAAAACCAATTGCTACTGATAGTGGCGGATGTGCACATAAACTTGAG	762	
Db	1309	TTAAGTCCAAAATCAAAACCAATTGCTACTGATAGTGGCGGATGTGCACATAAACTTGAG	1368	
Qy	763	AAAGTGACTTACTTAAGGGCTATTCAAGAACATTGATCGCTACGTCCACAGTAACGAC	822	

1369 AAAGCTGACTTAAAGGCTATTCAAGAACAAATTCATCGCTAAAGCTCCACAGTAACGAC 1428
823 GACTACTTTGAGTCAATTCATTTTGCAAGCGATGCAACATTAATCTGATGAAACGCGAAG 882
1429 GACTACTTTGAGTCAATTCATTTTGCAAGCGATGCAACATTAATCTGATGAAACGCGAAG 1488
883 GTCTACTTTGCTGACAAAGATGTTTCGGTAACCTTGCCGACCCAACTGTCGAAGAAATTT 942
1489 GTCTACTTTGCTGACAAAGATGTTTCGGTAACCTTGCCGACCCAACTGTCGAAGAAATTT 1548
943 TTGCTAAGCGGACATGTCGCGGTAGACCATATAAAGAAAAAACAATAACAAACCAAGCG 1002
1549 TTGCTAAGCGGACATGTCGCGGTAGACCATATAAAGAAAAAACAATAACAAACCAAGCG 1608
1003 AAATCTGTTGATGTGAATATATCTGATGACATGTTTATCTCCCTTAAACCCCTGATGACGATTC 1062
1609 AAATCTGTTGATGTGAATATATCTGATGACATGTTTATCTCCCTTAAACCCCTGATGACGATTC 1668
1063 AGACCAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 1122
1669 AGACCAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 1728
1123 TCTCAAGAAATTAAGTCAAGCAACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACG 1182
1729 TCTCAAGAAATTAAGTCAAGCAACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACG 1788
1183 ATTATGAACGATGATCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTA 1242
1789 ATTATGAACGATGATCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTA 1848
1243 CCAATGGATCAAGAGTTTACTTACCGTGTAAATAACCGGAACCAAGCTTATAGGATCAAT 1302
1849 CCAATGGATCAAGAGTTTACTTACCGTGTAAATAACCGGAACCAAGCTTATAGGATCAAT 1908
1303 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTAC 1362
1909 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGATCTCTGAGAAATATTAC 1968
1363 GTCCCTAAAAAAGGGGAAAGCGTATGATCCCTTTGATCGCACTCACTTGAACCTGTTTC 1422
1969 GTCCCTAAAAAAGGGGAAAGCGTATGATCCCTTTGATCGCACTCACTTGAACCTGTTTC 2028
1423 ACCATCAATAGCTTGTGATGATACCAACGAATTCCTAAAAAGTGAGAGCTCTTTAACA 1482
2029 ACCATCAATAGCTTGTGATGATACCAACGAATTCCTAAAAAGTGAGAGCTCTTTAACA 2088
1483 GCTAGCGAAGCTAATCTAGACTTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTA 1542
2089 GCTAGCGAAGCTAATCTAGACTTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTA 2148
1543 CTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAG 1602
2149 CTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGAGTAGAG 2208
1603 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGGAAGCGAACCAGGA 1662
2209 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGGAAGCGAACCAGGA 2268
1663 GAGAATGCTAGCTATCATTTAGCTTATGATTAAGATCGTTTATACCGAAGAACCGAGAA 1722
2269 GAGAATGCTAGCTATCATTTAGCTTATGATTAAGATCGTTTATACCGAAGAACCGAGAA 2328
1723 GTTTACAGTACTCGGTTTATACAGGACACCTTATACCTGATAACCCCTAACGACAAA 1779
2329 GTTTACAGTACTCGGTTTATACAGGACACCTTATACCTGATAACCCCTAACGACAAA 2385

RESULT 2
STRSKC STRSKC 2568 bp DNA linear BCT 26-APR-1993
LOCUS Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
DEFINITION
ACCESSION K02986

K02986.1 GI:153808
streptokinase.
Streptococcus dysgalactiae subsp. equisimilis
Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2568)
Malke, H., Roe, B. and Ferretti, J. J.
Nucleotide sequence of the streptokinase gene from Streptococcus
equisimilis H46A
Gene 34 (2-3), 357-362 (1985)
2989113
Original source text: S.equisimilis (strain H46A) DNA, clone pMF5.
Draft entry and hard copy of sequence for [1] kindly provided by
J.J.Ferretti, 03-SEP-1985.
The -35 and -10 regions are located at positions 760-765 and
781-786 respectively and an SD sequence at 808-813. Downstream
from the coding region inverted repeats (positions 2176-2190 and
2203-2217) are thought to function as transcription terminators.
The nucleotide sequence of skc does not support the hypothesis that
the gene has evolved by duplication and fusion, as suggested by
internal two-fold AA homologies of its product.
Location/Qualifiers
1. .2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
794. .>2141
/product="skc mRNA"
819. .2141
/notes="prestreptokinase"
/codon_start=11
/transl_table=11
/protein_id="AAA26974.1"
/db_xref="GI:153809"
translations="MKNYLSFGMFAALLFALTFTGTVNSVOAIGPEWLLDRPVSNNLSOL
VVSAGTVSGTNODISLKEPEIDLTSRPAHGKTEGLSPKSPKPTADSGAMSHKLEK
ADLLKAIQQLIANVNSNDYFEVIDFASDATITDRNGKYVFADKDSVLTPTQVQE
FLLSGHRVRYPEKEPIQNAKSVDEYTVPLNDDDFRGLDKLTKLLKTLAIGD
TITSQBLAAQASILNKHNPYTIYERDSSIVQTHNDNIDFRTILPMDQEFYRYVKNRQ
AVRINKSGLINEINNTDLISEKYVLLKKEKPYDPDFRSHLKLFTIKYVDVDTNELL
KSQLLTASRNLDRLDYPDKAKLLYNNLDAFGIMDYTLTGKVEDNDDHTNRIT
VVMGKEPEGENASYHLAYDKDRYTESEREVSYLYRTGTPTIPDNPNDK"
819. .896
sig_peptide
/notes="streptokinase signal peptide"
897. .2138
mat_peptide
/product="streptokinase"
ORIGIN 5 bp upstream of PstI site.
Query Match 69.8%; Score 1244.4; DB 1; Length 2568;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 537 TATTGCTGGACCTGAGTGGCTGTAGACCTGCATCTGTCAACAACAGCCAATTGGTTGT 596
DB 896 TATTGCTGGACCTGAGTGGCTGTAGACCTGCATCTGTCAACAACAGCCAATTGGTTGT 955
QY 597 TAGCGTTGCTGGTACTGTTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 656
DB 956 TAGCGTTGCTGGTACTGTTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 1015
QY 657 CGATCTAATCATCAGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATC 716
DB 1016 CGATCTAATCATCAGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATC 1075
QY 717 AAAACCATTTGCTACTGATAGTCGCGATGTCAATAAACTTGAGAAAGCTGACTTACT 776
DB 1076 AAAACCATTTGCTACTGATAGTCGCGATGTCAATAAACTTGAGAAAGCTGACTTACT 1135
QY 777 AAAAGGCTATTCAAGAACAAATTTGATCGCTAACCTGCACAGTAACACGACTACTTTGAGGT 836
DB 1136 AAAAGGCTATTCAAGAACAAATTTGATCGCTAACCTGCACAGTAACACGACTACTTTGAGGT 1195

QY	837	CATTGATTTTGCAGCGATGCAACCAATTA	CTGATCGAAACGGCAAGTCTACTTTGCTGA	896
Db	1196	CATTGATTTTGCAGCGATGCAACCAATTA	CTGATCGAAACGGCAAGTCTACTTTGCTGA	1255
QY	897	CAAGATGTTGCGTAACTTGGCGCCCAACCTG	TCCAAAGATTTTGGCTTAAGCGACA	956
Db	1256	CAAGATGTTGCGTAACTTGGCGCCCAACCTG	TCCAAAGATTTTGGCTTAAGCGACA	1315
QY	957	TGTGCGCGTTAGACCATATAAAGAAAAA	CCAAATACAAACCAACCGGAAATCTGTTGATGT	1016
Db	1316	TGTGCGCGTTAGACCATATAAAGAAAAA	CCAAATACAAACCGGAAATCTGTTGATGT	1375
QY	1017	GGAATATCTGTACAGTTTACTCCCTTAAAC	CCCTGATGACGATTTTCAGACAGCTCTCAA	1076
Db	1376	GGAATATCTGTACAGTTTACTCCCTTAAAC	CCCTGATGACGATTTTCAGACAGCTCTCAA	1435
QY	1077	AGTACTAGCTATTGAAACACATAGCTAT	CGGTGACACCATCATCTCAAGATTTACT	1136
Db	1436	AGTACTAGCTATTGAAACACATAGCTAT	CGGTGACACCATCATCTCAAGATTTACT	1495
QY	1137	AGCTCAAGCACAAAGCAATTTTAAACAA	AAAAACCAACCGGCTTATACGATTTATGAACGTGA	1196
Db	1496	AGCTCAAGCACAAAGCAATTTTAAACAA	AAAAACCAACCGGCTTATACGATTTATGAACGTGA	1555
QY	1197	CTCTCAATCGTCACTCATGACAATGACA	TTTTCCGTTACGATTTTACCAATGGATCAAGA	1256
Db	1556	CTCTCAATCGTCACTCATGACAATGACA	TTTTCCGTTACGATTTTACCAATGGATCAAGA	1615
QY	1257	GTTTACTTACCGTGTAAATCGGAA	CAAGCTTATAGGATCAATAAAAATCTGCTCT	1316
Db	1616	GTTTACTTACCGTGTAAATCGGAA	CAAGCTTATAGGATCAATAAAAATCTGCTCT	1675
QY	1317	GAATGAAGAAATAAACAACACTGACCT	GTCTCTGAGAAATATTTACCTCTTAAAAAAGG	1376
Db	1676	GAATGAAGAAATAAACAACACTGACCT	GTCTCTGAGAAATATTTACCTCTTAAAAAAGG	1735
QY	1377	GGAAAGCGGTATGATCCCTTTGATCG	CAGTCACTTGAACCTGTTTACCAATCAATACGT	1436
Db	1736	GGAAAGCGGTATGATCCCTTTGATCG	CAGTCACTTGAACCTGTTTACCAATCAATACGT	1795
QY	1437	TGATGTCGATACCAACGAATTTGTA	AAAAAGTAGAGCAGCTCTTAAACAGCTAGCGAACGTAA	1496
Db	1796	TGATGTCGATACCAACGAATTTGTA	AAAAAGTAGAGCAGCTCTTAAACAGCTAGCGAACGTAA	1855
QY	1497	CTTAGACTTCAGAGATTTATACGAT	CTCTGATAGGCTTAAACTACTCTACACAAATCT	1556
Db	1856	CTTAGACTTCAGAGATTTATACGAT	CTCTGATAGGCTTAAACTACTCTACACAAATCT	1915
QY	1557	CGATGCTTTTGGTATTATGGACTAT	ACCTTAACTGGAAAGTAGAGGATTAATCACGATGA	1616
Db	1916	CGATGCTTTTGGTATTATGGACTAT	ACCTTAACTGGAAAGTAGAGGATTAATCACGATGA	1975
QY	1617	CACCAACCGTATCATACCGTTTAT	TATGGCAAGCAGCCCGAGGAGAGATGCTAGCTA	1676
Db	1976	CACCAACCGTATCATACCGTTTAT	TATGGCAAGCAGCCCGAGGAGAGATGCTAGCTA	2035
QY	1677	TCATTAGCCTATGATAAGATCGTTA	TACCGAAGAAAGACGAGATTTACAGTACCT	1736
Db	2036	TCATTAGCCTATGATAAGATCGTTA	TACCGAAGAAAGACGAGATTTACAGTACCT	2095
QY	1737	GGTGTATACAGGACACCTTATACCT	GATACCCCTTAACGACAAATAA	1782
Db	2096	GGTGTATACAGGACACCTTATACCT	GATACCCCTTAACGACAAATAA	2141

RESULT 3
A04926
LOCUS A04926 2568 bp DNA linear PAT 15-JUL-1993
DEFINITION S.equisimilis skc gene for streptokinase.
ACCESSION A04926
VERSION A04926.1 GI:412219
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis

ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2568)
AUTHORS Ferretti,J.J. and Malke,H.
TITLE Streptokinase-coding recombinant vectors
JOURNAL Patent: EP 0151337-A 1 14-AUG-1985;
THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
FEATURES Location/Qualifiers
source 1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"

ORIGIN

Query Match	69.8%;	Score 1244.4;	DB 6;	Length 2568;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1245;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	537	TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACACGCCAATTTGGTTGT	596	
Db	896	TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACACGCCAATTTAGTTGT	955	
QY	597	TAGCGTTGCTGCTACTGTTGAGGGGACGGAATCAAGACATTAAGTCTTAAATTTTTTGAAT	656	
Db	956	TAGCGTTGCTGCTACTGTTGAGGGGACGGAATCAAGACATTAAGTCTTAAATTTTTTGAAT	1015	
QY	657	CGATCTAAATCATCAGACCTGCTCATGAGGAAAGACAGACGCAAGCTTTAAGTCCAAATC	716	
Db	1016	CGATCTAAATCATCAGACCTGCTCATGAGGAAAGACAGACGCAAGCTTTAAGTCCAAATC	1075	
QY	717	AAAACCATTTGCTACTGATGTCGGCGATGTCACATAAATCTTGAGAAAGCTGACTTACT	776	
Db	1076	AAAACCATTTGCTACTGATGTCGGCGATGTCACATAAATCTTGAGAAAGCTGACTTACT	1135	
QY	777	AAAGCTATTCAAGAAACAATTTGATCGCTAACGTCCACAGTAAACGACGACTACTTTTGAGGT	836	
Db	1136	AAAGCTATTCAAGAAACAATTTGATCGCTAACGTCCACAGTAAACGACGACTACTTTTGAGGT	1195	
QY	837	CATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGA	896	
Db	1196	CATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTCTACTTTGCTGA	1255	
QY	897	CAAAAGATGTTGCGTAACTTTGCGGACCCAAACCTGTGTCAGAAATTTTTCGTAAGCGACA	956	
Db	1256	CAAAAGATGTTGCGTAACTTTGCGGACCCAAACCTGTGTCAGAAATTTTTCGTAAGCGACA	1315	
QY	957	TGTGCGGTTAGACCATATAAAGAAAAA	CCAAATACAAACCAACCGGAAATCTGTTGATGT	1016
Db	1316	TGTGCGGTTAGACCATATAAAGAAAAA	CCAAATACAAACCAACCGGAAATCTGTTGATGT	1375
QY	1017	GGAATATCTGTACAGTTTACTCCCTTAAAC	CCCTGATGACGATTTTACAGACAGCTCTCAA	1076
Db	1376	GGAATATCTGTACAGTTTACTCCCTTAAAC	CCCTGATGACGATTTTACAGACAGCTCTCAA	1435
QY	1077	AGATACTAAGCTATTTGAAAAACATAGCTATCGGTGACACCATCAATCTCAAGAAATTTACT	1136	
Db	1436	AGATACTAAGCTATTTGAAAAACATAGCTATCGGTGACACCATCAATCTCAAGAAATTTACT	1495	
QY	1137	AGCTCAAGCACAAAGCAATTTTAAACAA	AAAAACCAACCGGCTTATACGATTTATGAACGTGA	1196
Db	1496	AGCTCAAGCACAAAGCAATTTTAAACAA	AAAAACCAACCGGCTTATACGATTTATGAACGTGA	1555
QY	1197	CTCTCAATCGTCACTCATGACAATGACA	TTTTCCGTTACGATTTTACCAATGGATCAAGA	1256
Db	1556	CTCTCAATCGTCACTCATGACAATGACA	TTTTCCGTTACGATTTTACCAATGGATCAAGA	1615
QY	1257	GTTTACTTACCGTGTAAATCGGAA	CAAGCTTTATAGGATCAATAAAAATCTGCTCT	1316
Db	1616	GTTTACTTACCGTGTAAATCGGAA	CAAGCTTTATAGGATCAATAAAAATCTGCTCT	1675
QY	1317	GAATGAAGAAATAAACAACACTGACCT	GTCTCTGAGAAATATTTACCTCTTAAAAAAGG	1376

CDS	complement (3262.. 4110) /gene="lrp" /codon_start=1 /transl_table=11 /product="leucine rich protein" /protein_id="CAA51350.1" /db_xref="GI:407878" /db_xref="GOA:Q54087" /db_xref="InterPro:IPR008931" /db_xref="UniProt/Swiss-Prot:Q54087" /translation="MELKDYFPEMVQPHPLGDKWVSVEGQYVHFPPKSCISEKER LLEVLGQYELVQLPGSQWRQLLDHQNPQLFETSFYILNHQOVLPAIDLVELIQ QMAGLEVILPISITTFALCRQATSIKVLRSLEGLPTLESDFGLALTFVGNAMWYQ VAAGTREFEECQLLTAYLKQKSGKGLITFAEVLMLWSILSHQSFPALTRQPHOFLN PQSDMADVVAHWLSEHNLVQTAQRLYIHRNSLQYKLDKFAQQSGSLHLKQLDLDLAPAY LFLKKY"	
misc_feature	complement (3337.. 3423) /gene="lrp" /note="leucine zipper motif"	Query Match 69.8%; Score 1244.4; DB 1; Length 8931; Best Local Similarity 99.9%; Pred. No. 0; Matches 1245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
RBS	complement (4117.. 4120) /gene="lrp"	
-10_signal	complement (4141.. 4146) /gene="lrp"	Qy 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCCCAATTGGTTGT 596
misc_feature	complement (4148.. 4149) /gene="lrp"	Db 4516 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCCCAATTAGTTGT 4575
misc_feature	complement (4172.. 4368) /note="TG motif"	Qy 597 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 656
gene	4392.. 5837 /gene="skc"	Db 4576 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 4635
misc_feature	4392.. 4393 /gene="skc"	Qy 657 CGATCTAACATCACGACCTGCTCATGAGGAAAGACAGCAAGCTTTAAGTCCAAATC 716
-10_signal	4395.. 4400 /note="TG motif"	Db 4636 CGATCTAACATCACGACCTGCTCATGAGGAAAGACAGCAAGCTTTAAGTCCAAATC 4695
-10_signal	4401.. 4406 /gene="skc"	Qy 717 AAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTACT 776
RBS	4428.. 4431 /note="alternative"	Db 4696 AAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTACT 4755
CDS	4439.. 5761 /gene="skc" /codon_start=1 /transl_table=11 /product="streptokinase" /protein_id="CAA51351.1" /db_xref="GI:407879" /db_xref="GOA:P00779" /db_xref="InterPro:IPR004093" /db_xref="InterPro:IPR008124" /db_xref="UniProt/Swiss-Prot:P00779"	Qy 777 AAAGGCTATTCAAGAACAAATTTGATCGCTAACGTCCACAGTAAACACGACTACTTTTGAGGT 836
sig_peptide	4439.. 4516 /gene="skc"	Db 4756 AAAGGCTATTCAAGAACAAATTTGATCGCTAACGTCCACAGTAAACGACGACTACTTTTGAGGT 4815
terminator	5796.. 5837 /gene="skc"	Qy 837 CATTGATTTTGCAGCGATGCAACCATTTACTGTAGTGGAAACCGCAAGGCTTACTTTGCTGA 896
CDS	complement (5859.. 6302) /note="unnamed protein product; ORF1" /codon_start=1 /transl_table=11 /protein_id="CAA51352.1" /db_xref="GI:407880" /db_xref="GOA:Q54088" /db_xref="InterPro:IPR003732" /db_xref="UniProt/Swiss-Prot:Q54088"	Db 4816 CATTGATTTTGCAGCGATGCAACCATTTACTGTAGTGGAAACCGCAAGGCTTACTTTGCTGA 4875
		Qy 897 CAAAGATGGTTCCGTAACCTTGGCGACCCAACTGTCCAGAAATTTTGTCAAGCGGACA 956
		Db 4876 CAAAGATGGTTCCGTAACCTTGGCGACCCAACTGTCCAGAAATTTTGTCAAGCGGACA 4935
		Qy 957 TGTGCGGTTAGACCATATATAAGAAAAACCAATAACAAACCAAGCGAAATCTGTTGATGT 1016
		Db 4936 TGTGCGGTTAGACCATATATAAGAAAAACCAATAACAAACCAAGCGAAATCTGTTGATGT 4995
		Qy 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAGACGAGCTCTCAA 1076
		Db 4996 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAGACGAGCTCTCAA 5055
		Qy 1077 AGATATTAAGCTATTGAAAAACATAGTCTATCGGTGACACCATCATCTCAAGAATTTACT 1136
		Db 5056 AGATATTAAGCTATTGAAAAACATAGTCTATCGGTGACACCATCATCTCAAGAATTTACT 5115
		Qy 1137 AGCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACTGA 1196
		Db 5116 AGCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACTGA 5175
		Qy 1197 CTCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGATCAAGA 1256
		Db 5176 CTCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGATCAAGA 5235
		Qy 1257 GTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 1316
		Db 5236 GTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 5295
		Qy 1317 GAATGAGAAATAAACAACACTGATCTCTGAGAAATATATACGTCCTTAAAAAAGG 1376

```

Db      5296  GAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAGG 5355
Qy      1377  GGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGCTTCCACCATCAATACGT 1436
Db      5356  GGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGCTTCCACCATCAATACGT 5415
Qy      1437  TGATGTCGATACCAACGAATTCGTAAGAAAGTGAAGAGCTCTTAAACAGCTAGCGAAGTAA 1496
Db      5416  TGATGTCGATACCAACGAATTCGTAAGAAAGTGAAGAGCTCTTAAACAGCTAGCGAAGTAA 5475
Qy      1497  CTTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGGCTTAACTACTCTACCAACATCT 1556
Db      5476  CTTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGGCTTAACTACTCTACCAACATCT 5535
Qy      1557  CGATGCTTTTGGTATATGCACTATACCTTAACTGGAAGAGTGAAGATTAATCACGATGA 1616
Db      5536  CGATGCTTTTGGTATATGCACTATACCTTAACTGGAAGAGTGAAGATTAATCACGATGA 5595
Qy      1617  CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGATGCTAGCTA 1676
Db      5596  CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGATGCTAGCTA 5655
Qy      1677  TCATTTAGCCTATGATAAAGATCGTTTATACCGAAGAAACCGAAGAGTTTACAGCTACCT 1736
Db      5656  TCATTTAGCCTATGATAAAGATCGTTTATACCGAAGAAACCGAAGAGTTTACAGCTACCT 5715
Qy      1737  GCGTTATACAGGACACCTATACCTGATTAACCTTAACCTTAACCTTAACCTTAACCTTA 1782
Db      5716  GCGTTATACAGGACACCTATACCTGATTAACCTTAACCTTAACCTTAACCTTAACCTTA 5761

RESULT 5
AR144000
LOCUS      AR144000      1242 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6210667.
ACCESSION AR144000
VERSION    AR144000.1 GI:15105867
KEYWORDS   Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 1242)
AUTHORS     Reed,G.L.
TITLE        Bacterial fibrin-dependent plasminogen activator
JOURNAL      Patent: US 6210667-A 5 03-APR-2001;
FEATURES     Location/Qualifiers
             source
               1..1242
               /organism="unknown"
               /mol_type="unassigned DNA"

ORIGIN
Query Match      69.6%; Score 1240.4; DB 6; Length 1242;
Best Local Similarity 99.9%; Pred. No. 2,8e-313;
Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      538  ATTGCTGACCTGAGTGGCTGCTAGACCGTCAATCTGTCAACAAACAGCCAAATTTGGTGT 597
Db      1  ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCCAAATTTAGTGT 60
Qy      598  ACGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
Db      61  ACGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy      658  GATCTAAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAGGCTTAAAGTCCAAATCA 717
Db      121  GATCTAAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAGGCTTAAAGTCCAAATCA 180
Qy      718  AAACATTTGCTACTGATAGTGGCGGATGTCAATATACTGAGAAAGCTGACTACTA 777
Db      181  AAACATTTGCTACTGATAGTGGCGGATGTCAATATACTGAGAAAGCTGACTACTA 240
Qy      778  AAGGCTATTCAAGAACATTTGATCGCTACGTCACGTAACGACGACTACTTTGAGTTC 837

```

```

Db      241  AAGGCTATTCAAGAACAAATTTGATCGCTTAAACGCTCCACAGTAACGACGACTACTTTGAGTTC 300
Qy      838  ATTGATTTTCCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 897
Db      301  ATTGATTTTCCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 360
Qy      898  AAGATGTTTCCGGTAAACCTTTGCCGACCAACCTCTCCAAAGAAATTTTGTCTAAGCGGACAT 957
Db      361  AAGATGTTTCCGGTAAACCTTTGCCGACCAACCTCTCCAAAGAAATTTTGTCTAAGCGGACAT 420
Qy      958  GTGCGGCTTAGAGCCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 1017
Db      421  GTGCGGCTTAGAGCCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 480
Qy      1018  GAATATATCTGACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCTCAA 1077
Db      481  GAATATATCTGACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAGACGAGGCTCTCAA 540
Qy      1078  GATATTAAGCTTATTTGAAAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTA 1137
Db      541  GATATTAAGCTTATTTGAAAAACACTAGCTATCGGTGACACCATCATCTCTCAAGAAATTA 600
Qy      1138  GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTTATGAACGCTGAC 1197
Db      601  GCTCAAGCACAAGCAATTTTAAACAAAAACCAAGGCTATACGATTTTATGAACGCTGAC 660
Qy      1198  TCCTCAATGCTCACTCATGACCAATGATTTTCCGTAGAGATTTTACCAATGATCAAGAG 1257
Db      661  TCCTCAATGCTCACTCATGACCAATGATTTTCCGTAGAGATTTTACCAATGATCAAGAG 720
Qy      1258  TTTTACTTACCGTGTTTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db      721  TTTTACTTACCGTGTTTAAATAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 780
Qy      1318  AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAAGGG 1377
Db      781  AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAACGCTCTTAAAAAAGGG 840
Qy      1378  GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACATCAATACGTT 1437
Db      841  GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTCAACATCAATACGTT 900
Qy      1438  GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAAC 1497
Db      901  GATGTCGATACCAACGAATTTGCTTAAAGTGAGCAGCTCTTAAACAGCTAGCGAAGTAAAC 960
Qy      1498  TTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGGCTAAACTTACTCTACCAATCTC 1557
Db      961  TTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGGCTAAACTTACTCTACCAATCTC 1020
Qy      1558  GATGCTTTTGGTATTTAGGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGAC 1617
Db      1021  GATGCTTTTGGTATTTAGGACTATACCTTAACTGGAAGAGTAGAGGATAATCACGATGAC 1080
Qy      1618  ACCAAGCTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1677
Db      1081  ACCAAGCTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1140
Qy      1678  CATTTAGCCTTATGATAAAGATCGTTTATACCGAAGAAACGAGAAAGTTTACAGCTACCTG 1737
Db      1141  CATTTAGCCTTATGATAAAGATCGTTTATACCGAAGAAACGAGAAAGTTTACAGCTACCTG 1200
Qy      1738  CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAA 1779
Db      1201  CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAA 1242

RESULT 6
E00522
LOCUS      E00522      2568 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION DNA fragment of plasmid PMF1 into which DNA encoding streptokinase is inserted.

```

E00522
 E00522.1 GI:2168801
 JP 1985237995-A/1.
 Streptococcus dysgalactiae subsp. equisimilis
 Streptococcus dysgalactiae subsp. equisimilis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 1. (bases 1 to 2568)
 JIYOSU, J.F. and HORUSUTO, M.
 RECOMBINATION VECTOR FOR PRODUCING STREPTOKINASE
 Patent: JP 1985237995-A 1 26-NOV-1985;
 PHILLIPS PETROLEUM CO
 OS Streptococcus equisimilis
 PN JP 1985237995-A/1
 PD 26-NOV-1985
 PF 09-OCT-1984 JP 1984212403
 PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PI
 JIYOSU JIEI FURETSUTSUI, HORUSUTO MARUKO
 PC C12N15/00, C12N1/20, C12N9/70, (C12N15/00, C12R1:46), (C12N1/20, PC
 C12R1:19),
 PC (C12N9/70, C12R1:19);
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: strains: Streptococcus equisimilis H46A; CC *source:
 library: streptococcus equisimilis H46 library; CC *source:
 clone: lambda L47 skc clone;
 FH key Location/Qualifiers
 FT 5'UTR 1..818
 FT sig_peptide 819..896
 FT /product= 'streptokinase signal peptide' FT
 mat_peptide 897..2138
 FT /product= 'streptokinase'
 FT CDS 819..2143
 FT /product= 'streptokinase'
 FT 3'UTR 2142..2569.
 FEATURES
 source
 1..2568
 Location/Qualifiers
 /organism= 'Streptococcus dysgalactiae subsp. equisimilis'
 /mol_type= 'genomic DNA'
 /db_xref= 'taxon:119602'
 ORIGIN
 Query Match 69.6%; Score 1239.6; DB 6; Length 2568;
 Best Local Similarity 99.7%; Pred. No. 5.6e-313;
 Matches 1242; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 537 TATTGCTGAGCTGAGTGGCTGCTAGACCGTCCATCTGTCACAAACAGCGCAATTGGTTGT 596
 896 TATTGCTGAGCTGAGTGGCTGCTAGACCGTCCATCTGTCACAAACAGCGCAATTGGTTGT 955
 597 TAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTCAAAT 656
 956 TAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTCAAAT 1015
 657 CGATCTAATCAAGACCTGCTCATGAGGAAAGACAGAGCGGCTTAAGTCCAAATC 716
 1016 CGATCTAATCAAGACCTGCTCATGAGGAAAGACAGAGCGGCTTAAGTCCAAATC 1075
 717 AAAACCAATTCCTACTGATAGTGGCGGATGTCACATAAATTCGAAAGCTGACTTACT 776
 1076 AAAACCAATTCCTACTGATAGTGGCGGATGTCACATAAATTCGAAAGCTGACTTACT 1135
 777 AAAGCTATTCAAGAACCAATTCATCGCTAACCTCCACAGTACGAGCTACTTTGAGGT 836
 1136 AAAGCTATTCAAGAACCAATTCATCGCTAACCTCCACAGTACGAGCTACTTTGAGGT 1195
 837 CATTGATTTGCAAGCGATGCAACCAATTCATGATCGAAACGGCAAGGTCTACTTTGCTGA 896
 1196 CATTGATTTGCAAGCGATGCAACCAATTCATGATCGAAACGGCAAGGTCTACTTTGCTGA 1255

QY 897 CAAAGATGGTTCGGTAAACCTTGCAGACCCAAACCTGTCCAGAAATTTTGTCAAGCGACA 956
 DB 1256 CAAAGATGGTTCGGTAAACCTTGCAGACCCAAACCTGTCCAGAAATTTTGTCAAGCGACA 1315
 QY 957 TGTGCGGTTAGACCATATAAGAAAAACCAATCAAAACCAAGGAAATCTGTTGATGT 1016
 DB 1316 TGTGCGGTTAGACCATATAAGAAAAACCAATCAAAACCAAGGAAATCTGTTGATGT 1375
 QY 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 1076
 DB 1376 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 1435
 QY 1077 AGATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCTCAAGAAATTACT 1136
 DB 1436 AGATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCTCAAGAAATTACT 1495
 QY 1137 AGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGA 1196
 DB 1496 AGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGA 1555
 QY 1197 CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1256
 DB 1556 CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1615
 QY 1257 GTTTACTTACCGTGTAAATAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGCTCT 1316
 DB 1616 GTTTACTTACCGTGTAAATAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGCTCT 1675
 QY 1317 GAATGAAGAAATAAACCAACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1376
 DB 1676 GAATGAAGAAATAAACCAACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1735
 QY 1377 GGAAGAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCCACCATCAATACGT 1436
 DB 1736 GGAAGAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCCACCATCAATACGT 1795
 QY 1437 TGATGTCGATACCAACGAATTTGCTTAAAGAGTGACAGCTCTTAAACAGCTAGCGAACTAA 1496
 DB 1796 TGATGTCGATACCAACGAATTTGCTTAAAGAGTGACAGCTCTTAAACAGCTAGCGAACTAA 1855
 QY 1497 CTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTTAACTCTCTACAACTCT 1556
 DB 1856 CTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTTAACTCTCTACAACTCT 1915
 QY 1557 CGATGCTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGATTAATCAAGATGA 1616
 DB 1916 CGATGCTTTGGTATTATGGAATATACCTTAACTGGAAGAGTAGAGATTAATCAAGATGA 1975
 QY 1617 CACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1676
 DB 1976 CACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 2035
 QY 1677 TCATTTAGCTTATGATAAAGATCGTTATACCGAAGAGAAACGAGAACTTTACAGCTACCT 1736
 DB 2036 TCATTTAGCTTATGATAAAGATCGTTATACCGAAGAGAAACGAGAACTTTACAGCTACCT 2095
 QY 1737 GCCTTATACAGGACACCTTATACCTGATAACCTTAACCGACAAATAA 1782
 DB 2096 GCCTTATACAGGACACCTTATACCTGATAACCTTAACCGACAAATAA 2141
 RESULT 7
 105204
 LOCUS
 DEFINITION
 Sequence 5 from Patent EP 0248227.
 105204
 ACCESSION
 105204.1 GI:591209
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 Unclassified.
 1 (bases 1 to 1401)
 REFERENCE
 Hagenson, M.J. and Stroman, D.W.

TITLE Yeast production of streptokinase
JOURNAL Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES Location/Qualifiers
source 1. .1401
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 69.4%; Score 1237; DB 6; Length 1401;
Best Local Similarity 99.6%; Pred. No. 3.1e-312;
Matches 1240; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGT 597
Db 10 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTAGTTGT 69
Qy 598 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 657
Db 70 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 129
Qy 658 GATCTAAACATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 717
Db 130 GATCTAAACATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAAGTCCAAATCA 189
Qy 718 AAACCATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTGAGAAAGCTGACTTACTA 777
Db 190 AAACCATTTGCTACTGATAGTGGCGGATGTCTACATAAACTTGAGAAAGCTGACTTACTA 249
Qy 778 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCACAGTAAACGACACTTGTAGGTC 837
Db 250 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCACAGTAAACGACACTTGTAGGTC 309
Qy 838 ATTGATTTTGAAGCGATCAACCTTACTGATCGAAGCGGACGCTCTACTTTGCTGAC 897
Db 310 ATTGATTTTGAAGCGATCAACCTTACTGATCGAAGCGGACGCTCTACTTTGCTGAC 369
Qy 898 AAAGATGTTTGGTAACTTGTCCGACCCCACTGTCCAAAGAAATTTTGTCTAAGCGGACAT 957
Db 370 AAAGATGTTTGGTAACTTGTCCGACCCCACTGTCCAAAGAAATTTTGTCTAAGCGGACAT 429
Qy 958 GTGGCGGTTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 430 GTGGCGGTTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATCTGTTGATGTG 489
Qy 1018 GAATATCTGTACGTTTACTCCCTTAAACCTGTATGACGATTTACAGCAGTCTCAAA 1077
Db 490 GAATATCTGTACGTTTACTCCCTTAAACCTGTATGACGATTTACAGCAGTCTCAAA 549
Qy 1078 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 1137
Db 550 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 609
Qy 1138 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
Db 610 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 669
Qy 1198 TCCTCAATCTGTACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 1257
Db 670 TCCTCAATCTGTACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 729
Qy 1258 TTTACTTACCGTGTAAAACCGGAAACAGCTTTATAGGATCAATAAAAACTCGTCTG 1317
Db 730 TTTACTTACCGTGTAAAACCGGAAACAGCTTTATAGGATCAATAAAAACTCGTCTG 789
Qy 1318 AATGAAGAAATAAACAACACACTGACCTGTCTCTGAGAAATATTTACGTCCTTAAAAAGGG 1377
Db 790 AATGAAGAAATAAACAACACACTGACCTGTCTCTGAGAAATATTTACGTCCTTAAAAAGGG 849
Qy 1378 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTGTCAACATCAAAATACGTT 1437
Db 850 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTGTCAACATCAAAATACGTT 909
Qy 1438 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1497

Db 910 GATGTCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 969
Qy 1498 TTAGACTTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTTCTTACAAACAATCTC 1557
Db 970 TTAGACTTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTTCTTACAAACAATCTC 1029
Qy 1558 GATGCTTTTGGTATTATGGAATATACCTTAACCTGAAAAAGTAGAGGATAATCAGCATGAC 1617
Db 1030 GATGCTTTTGGTATTATGGAATATACCTTAACCTGAAAAAGTAGAGGATAATCAGCATGAC 1089
Qy 1618 ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGAGAGAATGTAGCTAT 1677
Db 1090 ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGAGAGAATGTAGCTAT 1149
Qy 1678 CATTTAGCCTTATGATTAAGATCGTTATACCGAAGAAAGACGAGAGTTTACAGCTACTG 1737
Db 1150 CATTTAGCCTTATGATTAAGATCGTTATACCGAAGAAAGACGAGAGTTTACAGCTACTG 1209
Qy 1738 CGTTATACGGGACACCTATACCTGTAAACCTTAACGACAAATAA 1782
Db 1210 CGTTATACGGGACACCTATACCTGTAAACCTTAACGACAAATAA 1254
RESULT 8
LOCUS AR068768 2566 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 19 from patent US 5854049.
ACCESSION AR068768
VERSION AR068768.1 GI:6000975
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2566)
AUTHORS Reed,G.L.
TITLE Plasmin-resistant streptokinase
JOURNAL Patent: US 5854049-A 19 29-DEC-1998;
FEATURES Location/Qualifiers
source 1. .2566
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 69.0%; Score 1230.4; DB 6; Length 2566;
Best Local Similarity 99.8%; Pred. No. 1.7e-310;
Matches 1243; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAATTTGGTTGT 596
Db 896 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAATTTAGTTGT 955
Qy 597 TAGCGTTGCTGTGCTGTTTGGAGGGAGCAATCAAGACATTTAGTCTTAAATTTTGAAT 656
Db 956 TAGCGTTGCTGTGCTGTTTGGAGGGAGCAATCAAGACATTTAGTCTTAAATTTTGAAT 1015
Qy 657 CGATCTAAACATACGACCTGCTCATGAGGAAAGACAGACAGCGCTTAGTCCAAATC 716
Db 1016 CGATCTAAACATACGACCTGCTCAT - AGGAAAGACAGACAGCGCTTAGTCCAAATC 1073
Qy 717 AAAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACT 776
Db 1074 AAAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACT 1133
Qy 777 AAAGGCTATTTCAGAAACAAATTTGATCGCTAACGTCCTCAAGTAAACGACGACTCTTTGAGGT 836
Db 1134 AAAGGCTATTTCAGAAACAAATTTGATCGCTAACGTCCTCAAGTAAACGACGACTCTTTGAGGT 1193
Qy 837 CATTTGATTTTGAAGCGATGCAACCATTAATGATCGAAACGGCAAGGTCTACTTTGCTGA 896
Db 1194 CATTTGATTTTGAAGCGATGCAACCATTAATGATCGAAACGGCAAGGTCTACTTTGCTGA 1253
Qy 897 CAAGATGGTTCGGTAACTTGGCGGACCCCAACCTGTCCAAGAAATTTTGTCTAAGCGGACA 956

QY 1426 ATCAATACGTTGATGTCGATACCAACGAATTCCTAAAGAGTGGACGAGCTCTTTAACAGCT 1485
Db 1096 ATCAATACGTTGATGTCGATACCAACGAATTCCTAAAGAGTGGACGAGCTCTTTAACAGCT 1155
QY 1486 AGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAACTACTC 1545
Db 1156 AGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAACTACTC 1215
QY 1546 TACAACAACTCGATGCTTTTGGTATTTATGAGCTATACCTTTAACTGCGAAGTAGAGAT 1605
Db 1216 TACAACAACTCGATGCTTTTGGTATTTATGAGCTATACCTTTAACTGCGAAGTAGAGAT 1275
QY 1606 AATCAGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCACCAGGAGAGAG 1665
Db 1276 AATCAGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCACCAGGAGAGAG 1335
QY 1666 AATGCTAGCTATCAATTTAGCTTATGATTAAGATTCGTTATACCGAAGAACGAGAAGTT 1725
Db 1336 AATGCTAGCTATCAATTTAGCTTATGATTAAGATTCGTTATACCGAAGAACGAGAAGTT 1395
QY 1726 TACAGCTACCTGGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAATA 1782
Db 1396 TACAGCTACCTGGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAATA 1452

RESULT 10
113215
LOCUS 113215 1458 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 42 from patent US 5434073.
ACCESSION 113215
VERSION 113215.1 GI:910563
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1458)
AUTHORS Dawson, K., Hunter, M.G. and Czaplowski, L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 42 18-JUL-1995;
FEATURES
source
1. .1458
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 68.7%; Score 1223.4; DB 6; Length 1458;
Best Local Similarity 98.3%; Pred. No. 1.1e-308;
Matches 1236; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 526 ACCGATGTTGTTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGC 585
Db 196 ATCGAAGGTAGAAATTTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGC 255
QY 586 CAAFTGTTGTTAGCGTGTCTGCTAGTCTGTTAGGGGACGAATCAAGACATTTAGTCTTAAA 645
Db 256 CAATTTAGTTGTTAGCGTGTCTGCTAGTCTGTTAGGGGACGAATCAAGACATTTAGTCTTAAA 315
QY 646 TTTTGTGAAATCGATTAACATCAACGCTGCTCATCGAGGAAAGACAGAGCAAGGCTTA 705
Db 316 TTTTGTGAAATCGATTAACATCAACGCTGCTCATCGAGGAAAGACAGAGCAAGGCTTA 375
QY 706 AGTCCAAATCAAAACCAATTTGCTACTGATAGTGGGCGGATGTCACATTAACCTTGAGAAA 765
Db 376 AGTCCAAATCAAAACCAATTTGCTACTGATAGTGGGCGGATGTCACATTAACCTTGAGAAA 435
QY 766 GCTGACTTAAAGGCTATTCAAGAACCAATTCATCGTAACTGCCACGTAACGACGAC 825
Db 436 GCTGACTTAAAGGCTATTCAAGAACCAATTCATCGTAACTGCCACGTAACGACGAC 495
QY 826 TACTTTGAGGTCAATTTGATTTGCAAGCGATGCAACCACTTACTGATCGAAACGGCAGGTC 885
Db 496 TACTTTGAGGTCAATTTGATTTGCAAGCGATGCAACCACTTACTGATCGAAACGGCAGGTC 555

QY 886 TACTTTGCTGCAAAAGATGGTTCCGTTAACTTGCAGCCCAACCTGTCTCAAGAAATTTTGTG 945
Db 556 TACTTTGCTGCAAAAGATGGTTCCGTTAACTTGCAGCCCAACCTGTCTCAAGAAATTTTGTG 615
QY 946 CTAAGCGGACATGTGCGCGTTAGACCATATATAAGAAACCAATACAAAACCAAGCGGAAA 1005
Db 616 CTAAGCGGACATGTGCGCGTTAGACCATATATAAGAAACCAATACAAAACCAAGCGGAAA 675
QY 1006 TCTGTTGATGTGGAAATACCTGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTACGA 1065
Db 676 TCTGTTGATGTGGAAATACCTGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTACGA 735
QY 1066 CCAGGTCTCAAGAGATATAAGCTATTGAAAACACATAGCTATCGGTGACACCATCACATCT 1125
Db 736 CCAGGTCTCAAGAGATATAAGCTATTGAAAACACATAGCTATCGGTGACACCATCACATCT 795
QY 1126 CAAGAATTACTAGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGCGCTATACGATT 1185
Db 796 CAAGAATTACTAGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGCGCTATACGATT 855
QY 1186 TATGAAGTGTGACTCTCAATCGTCACTCATGACGAATGACATTTTCCGTACGATTTTACCA 1245
Db 856 TATGAAGTGTGACTCTCAATCGTCACTCATGACGAATGACATTTTCCGTACGATTTTACCA 915
QY 1246 ATGATCAAGAGTTTACTTTACCGGTATAAAATCGGAAACCAAGCTTATAGGATCAATAAA 1305
Db 916 ATGATCAAGAGTTTACTTTACCATGTCTAAATCGGGAACAGCTTATGATCANATAA 975
QY 1306 AAATCTGCTGTAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTTACGTC 1365
Db 976 AAATCTGCTGTAATGAAGAAATAAACAACCTGATCTCTGAGAAATATTTACGTC 1035
QY 1366 CTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAAC 1425
Db 1036 CTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAAC 1095
QY 1426 ATCAATATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGGAGAGCTCTTAAACAGCT 1485
Db 1096 ATCAATATACGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGGAGAGCTCTTAAACAGCT 1155
QY 1486 AGCGAAGCTAACTTTAGACTTTTACAGAGATTTATATACGATCCTCGTGAAGGCTTAACTACTC 1545
Db 1156 AGCGAAGCTAACTTTAGACTTTTACAGAGATTTATATACGATCCTCGTGAAGGCTTAACTACTC 1215
QY 1546 TACAACAACTCTCGATGCTTTTGGTATTTATGAGCTATACCTTAACTGAGAAAGTGAAGAT 1605
Db 1216 TACAACAACTCTCGATGCTTTTGGTATTTATGAGCTATACCTTAACTGAGAAAGTGAAGAT 1275
QY 1606 AATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCAGCCGAGGAGAG 1665
Db 1276 AATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCAGCCGAGGAGAG 1335
QY 1666 AATGCTAGCTATCAATTTAGCTTATGATTAAGATCGTTATACCGAAGAACGAGAGATT 1725
Db 1336 AATGCTAGCTATCAATTTAGCTTATGATTAAGATCGTTATACCGAAGAACGAGAGATT 1395
QY 1726 TACAGCTACCTGCTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAATAA 1782
Db 1396 TACAGCTACCTGCTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAATAA 1452

RESULT 11
S46536
LOCUS
DEFINITION SKC-2=streptokinase 1245 bp DNA linear BCT 08-MAY-1993
Genomic, 1245 nt].
ACCESSION S46536
VERSION S46536.1 GI:257196
KEYWORDS
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE 1 (bases 1 to 1245)
AUTHORS Estrada,M.P., Hernandez,L., Perez,A., Rodriguez,P., Serrano,R.,
Rubiera,R., Pedraza,A., Padron,G., Antuch,W., de la Fuente,J. et
al.
TITLE High level expression of streptokinase in *Escherichia coli*
JOURNAL Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)
PUBMED 1368792
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 115306] from the original journal article.
FEATURES
source
1..1245
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
/notes="group: C"
1..1245
/gene="SKC-2"
1..1245
/gene="SKC-2"
/notes="SKC-2"
/notes="streptokinase; Mature protein lacks initial Met"
/codon_start=1
/product="SKC-2"
/protein_id="AAC60418.1"
/db_xref="GI:257197"
/translation="IAGPEWLLDRPSVNSQLVSVAGTVEGTNQDISLKFPEIDLTS
RPAHGKTEGLSPKSPKFPATDSGAMPKLEKADLLKAIQELIANVHSDYDEVID
FASDATITDRNGKVPADKQSVTLPTQPVQFLLSGHVRVPYKEKPIQNOAKSVDV
EYVQFTPLNDDPRPLKDLTKLLKTLAIGDITTSOELLAQAQSIINKTHPGYIYE
ROSSIVTHNDIIPRTILPMDQETPHVNRQAEYINKSGLNEINNDLISIKYIV
LKGEKPPDFDRSHUKLFTIKYVDVNTNLSKSEQLLTASERNLDFRDLVDRDKAK
LYNNIDAFGIMDYITUTGKVEDNHDNTRNIITVYMKRPEGENASYHLAYDKORYTEE
EREVSYLRYTGTPIPDNPNDK"
ORIGIN
Query Match 58.6%; Score 1222.6; DB 1; Length 1245;
Best Local Similarity 98.9%; Pred. No. 1.8e-308;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAGCAATGCTTGT 597
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAGCAATGCTTGT 60
QY 598 AGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 657
DB 61 AGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 120
QY 658 GATCTAACTACGACCTGCTCTATGGAGGAAAGACAGACAGAGCTTAAAGTCCAAATCA 717
DB 121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGACAGAGCTTAAAGTCCAAATCA 180
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAGCTGACTTACTA 777
DB 181 AAACCATTTGCTACTGATAGTGGCGGATGCCATATAAATTGAAAAGCTGACTTACTA 240
QY 778 AAGGCTATTCAAGAACAAATTGATCGCTTAACGTCACAGTAAACGACACTACTTTGAGTC 837
DB 241 AAGGCTATTCAAGAACAAATTGATCGCTTAACGTCACAGTAAACGACACTACTTTGAGTC 300
QY 838 ATGATTTTGAAGCGATGCAACATTAATGATCGAAACGCAAGGCTACTTTGCTGAC 897
DB 301 ATGATTTTGAAGCGATGCAACATTAATGATCGAAACGCAAGGCTACTTTGCTGAC 360
QY 898 AAAGATGGTTCGGTACCTTGGCGACCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT 957
DB 361 AAAGATGGTTCGGTACCTTGGCGACCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT 420
QY 958 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATG 1017
DB 421 GTGCGGTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAATCTGTTGATG 480
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTAGCAGGATTTTCAGACAGGCTCTCAA 1077
DB 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTAGCAGGATTTTCAGACAGGCTCTCAA 540

QY 1078 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 1137
DB 541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 600
QY 1138 GCTCAAGCACAAGCAATTTTAAACAAACACCCAGGGCTATACGATTATGAACGTGAC 1197
DB 601 GCTCAAGCACAAGCAATTTTAAACAAACACCCAGGGCTATACGATTATGAACGTGAC 660
QY 1198 TCCTCAATCGTCACATCATGACAATGACATTTTCCGTACGATTTTACCAGTCAAGAG 1257
DB 661 TCCTCAATCGTCACATCATGACAATGACATTTTCCGTACGATTTTACCAGTCAAGAG 720
QY 1258 TTCTTACTCCGTGTTTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTCGTCTG 1317
DB 721 TTCTTACTCCGTGTTTAAATAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTCGTCTG 780
QY 1318 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATATACGCTTAAAAAAGG 1377
DB 781 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATATACGCTTAAAAAAGG 840
QY 1378 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTTGAAACTGTTCAACATCAATACGTT 1437
DB 841 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTTGAAACTGTTCAACATCAATACGTT 900
QY 1438 GATGTCGATACCAACGAATTTGCTTAAAGATGAGCAGCTCTTAAACAGTAGCGAATCAAC 1497
DB 901 GATGTCACACCAACGAATTTGCTTAAAGATGAGCAGCTCTTAAACAGTAGCGAATCAAC 960
QY 1498 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAGGCTTAAACTCTCTACAACTCTC 1557
DB 961 TTAGACTTCAGAGATTTTATACGATCCTCGTGATAGGCTTAAACTCTCTACAACTCTC 1020
QY 1558 GATGCTTTGTTATATCGACTATACCTTAATCTGAAAGATGAGGATTAATCAGATGAC 1617
DB 1021 GATGCTTTGTTATATCGACTATACCTTAATCTGAAAGATGAGGATTAATCAGATGAC 1080
QY 1618 ACCAACCGTATCATAACTGTTTATATGCGAAGCAGCCGAGGAGAGAGTCTAGCTAT 1677
DB 1081 ACCAACCGTATCATAACTGTTTATATGCGAAGCAGCCGAGGAGAGAGTCTAGCTAT 1140
QY 1678 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAAACGAGAAAGTTTACAGCTACCTG 1737
DB 1141 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAAACGAGAAAGTTTACAGCTACCTG 1200
QY 1738 CGTTATACAGGACCACTTATACCTGATTAACCTTAACGACAAATAA 1782
DB 1201 CGTTATACAGGACCACTTATACCTGATTAACCTTAACGACAAATAA 1245
RESULT 12
QY97820 7057 bp DNA linear PAT 20-APR-2004
LOCUS Sequence 26 from Patent WO2004029256.
DEFINITION CQ797820
ACCESSION CQ797820
VERSION CQ797820.1 GI:46426093
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS selman-houein Sosa,G., aguiar Cabeza,E., gonz lez Quintero,A.D.
and ramos gonz lez,O.
TITLE Vector for the production of transplastomic angiosperm plants
JOURNAL Patent: WO 2004029256-A 26 08-APR-2004;
CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
FEATURES
source
1..7057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Nucleotide
sequence of DNA fragment from the vector pVTPA-Estrep

between the rice atpB and tobacco rbcL borders."

ORIGIN		Query Match	Best Local Similarity	Mismatches	Conservative	0	Indels	0	Gaps	0
Qy	538	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATTTGGTTGT	597							
Db	2023	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATTTGGTTGT	2082							
Qy	598	AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAATC	657							
Db	2083	AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAATC	2142							
Qy	658	GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA	717							
Db	2143	GACCTAACATCAGCACTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA	2202							
Qy	718	AAACCAATTTGCTACTGATAGTGGCGGATCTCATATAAATCTGAGAAAGCTGACTACTA	777							
Db	2203	AAACCAATTTGCTACTGATAGTGGCGGATCTCATATAAATCTGAGAAAGCTGACTACTA	2262							
Qy	778	AAGGCTATTCAAGACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC	837							
Db	2263	AAGGCTATTCAAGACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC	2322							
Qy	838	ATTGATTTTGAAGCGATGCAACATTTAGTATGATGCGAAACCGGAGGCTACTTTGCTGAC	897							
Db	2323	ATTGATTTTGAAGCGATGCAACATTTAGTATGATGCGAAACCGGAGGCTACTTTGCTGAC	2382							
Qy	898	AAAGATGTTTCGGTACCTTGGCGACCCACCTGTCACAGAAATTTTGTAAAGCGGCAT	957							
Db	2383	AAAGATGTTTCGGTACCTTGGCGACCCACCTGTCACAGAAATTTTGTAAAGCGGCAT	2442							
Qy	958	GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	1017							
Db	2443	GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG	2502							
Qy	1018	GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCAAA	1077							
Db	2503	GAATATACGTGACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCAAA	2562							
Qy	1078	GATCTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA	1137							
Db	2563	GATCTAAGCTATTGAACAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA	2622							
Qy	1138	GCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC	1197							
Db	2623	GCTCAAGCACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTTATGAACGTGAC	2682							
Qy	1198	TCCTCAATCGTCACTCATGACATGACATTTTCCGTCGATGATTTTACCAATGGATCAAGAG	1257							
Db	2683	TCCTCAATCGTCACTCATGACATGACATTTTCCGTCGATGATTTTACCAATGGATCAAGAG	2742							
Qy	1258	TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	1317							
Db	2743	TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG	2802							
Qy	1318	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	1377							
Db	2803	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG	2862							
Qy	1378	GAAGACCGGTATGATCCCTTTTGAATCGCAGTCACTTGAAACCTGTTCCACCAATACGTT	1437							
Db	2863	GAAGACCGGTATGATCCCTTTTGAATCGCAGTCACTTGAAACCTGTTCCACCAATACGTT	2922							
Qy	1438	GATGTCGATACCAACGAATTTGCTTAAAGATGAGCAGCTTTTAAAGCTACGCAACGTAAC	1497							
Db	2923	GATGTCGATACCAACGAATTTGCTTAAAGATGAGCAGCTTTTAAAGCTACGCAACGTAAC	2982							
Qy	1498	TTAGACTTCAGAGATTTTACGATCTCTGATGATAGGCTTAACTCTCTACACCAATCTC	1557							
Db	2983	TTAGACTTCAGAGATTTTATACGATCTCTGATGATAGGCTTAACTCTCTACACCAATCTC	3042							
Qy	1558	GATGCTTTTGGTATTATGAGTACTTAACTCGAAGATAGAGGATAATCAGATGAC	1617							
Db	3043	GATGCTTTTGGTATTATGAGTACTTAACTCGAAGATAGAGGATAATCAGATGAC	3102							
Qy	1618	ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT	1677							
Db	3103	ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT	3162							
Qy	1678	CATTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAAGTTTACAGTACCTG	1737							
Db	3163	CATTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAAGTTTACAGTACCTG	3222							
Qy	1738	CGTTATACAGGACACCTATACCTGATTAACCTTAACGCAATTA	1782							
Db	3223	CGTTATACAGGACACCTATACCTGATTAACCTTAACGCAATTA	3267							
RESULT 13										
A20006										
LOCUS										
DEFINITION										
ACCESSION										
VERSION										
KEYWORDS										
SOURCE										
ORGANISM										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
FEATURES										
source										
1. 1335										
/organism="Streptococcus dysgalactiae subsp. equisimilis"										
/mol_type="unassigned DNA"										
/sub_species="equisimilis"										
/db_xref="taxon:119602"										
7. 1329										
/codon_start=1										
/transl_table=11										
/product="streptokinase"										
/protein_id="CAA01484.1"										
/db_xref="GI:1247580"										
/translation="MKNYLSFGMPALLFALTFTGTVNSVQAIAGPEWLLDRPSVNSQL										
VVSVAGTVEGTNODISLKFPEIDLTSPHGGKTEQGLSPKSPKPTDSDGAMPHLEK										
ADLLKALQEQLIANVHSNDYFEVIDPASDATITDRNGKYVYFADKDSVTLPTQVQS										
FLLSGHVRVPIKEKPIQONQKASVDVETVQTFINPDDDFRPGKDKTLKLLKLAIGD										
TITSQELLAQAQSLINKTHPGYTIYERDSSIVTHDNDI FRTILPMDOEFTYHVKRQ										
AYEINKSGLINEINNTDLISEKYYLVKGEKPYDFDRSHLKLFTI KYVDVNTNELL										
KSEOLLTASERNLDFRLYDPRDKALLNNLDAFGIMDYTLTKGVEDHDDTNRIT										
VYMKRPEGENASVHLAYDKDRYTEEBEREVYSYLRYGTPTPDNPDK"										
ORIGIN										
Query Match										
Best Local Similarity										
Matches 1230; Conservative										
0; Mismatches 16; Indels 0; Gaps 0;										
Qy	537	TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATTTGGTTGT	596							
Db	84	TATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATTTAGTTGT	143							
Qy	597	TAGCGTTGCTGTTAGTGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAAT	656							
Db	144	TAGCGTTGCTGTTAGTGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAAT	203							
Qy	657	CGATCTAACATCAGACCTGCTCATGAGGAGAAAGACAGCAAGCTTTAAGTCCAAATC	716							
Db	204	TGACCTAACATCAGACCTGCTCATGAGGAGAAAGACAGCAAGCTTTAAGTCCAAATC	263							
Qy	717	AAACCAATTTGCTACTGATGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTACT	776							

Db 264 AAAACCAATTTGCTACTGATAGTGGCGATGCCACATAAACTTGAAAAAGCTGACTTACT 323
Qy 777 AAAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTTGAGGT 836
Db 324 AAAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTTGAGGT 383
Qy 837 CATGTATTTGCAAGCGATGCAACATTTATGTATCGAAACGCGAAGGTCCTACTTTGCTGA 896
Db 384 CATGTATTTGCAAGCGATGCAACATTTATGTATCGAAACGCGAAGGTCCTACTTTGCTGA 443
Qy 897 CAAAGATGTTCCGTAACCTTTCGCGACCCACCTGTCGAGAAATTTTGCTAAGCGGACA 956
Db 444 CAAAGATGTTCCGTAACCTTTCGCGACCCACCTGTCGAGAAATTTTGCTAAGCGGACA 503
Qy 957 TGTGCGGCTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGT 1016
Db 504 TGTGCGGCTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGT 563
Qy 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 1076
Db 564 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 623
Qy 1077 AGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1136
Db 624 AGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 683
Qy 1137 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTA 1196
Db 684 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCATCCAGGCTATACGATTTATGAACGTA 743
Qy 1197 CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCA 1256
Db 744 CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCA 803
Qy 1257 GTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 1316
Db 804 GTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 863
Qy 1317 GAATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGG 1376
Db 864 GAATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGG 923
Qy 1377 GGAAGAGCGCTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTTACCACATCAATACGT 1436
Db 924 GGAAGAGCGCTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTTACCACATCAATACGT 983
Qy 1437 TGATGTCGATACCAACGAAATTTGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAA 1496
Db 984 TGATGTCGATACCAACGAAATTTGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAA 1043
Qy 1497 CTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTTAACTACTTACAAATCT 1556
Db 1044 CTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTTAACTACTTACAAATCT 1103
Qy 1557 CGATGCTTTTGGTATTATGCACTATACCTTAACTGGAAGAGTACGATTAATCAGTGA 1616
Db 1104 CGATGCTTTTGGTATTATGCACTATACCTTAACTGGAAGAGTACGATTAATCAGTGA 1163
Qy 1617 CACCAACCGCTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1676
Db 1164 CACCAACCGCTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1223
Qy 1677 TCATTTAGCTATGATAAGATCGTTATACGGAAGAGAACGAGAGTTTACAGTACCT 1736
Db 1224 TCATTTAGCTATGATAAGATCGTTATACGGAAGAGAACGAGAGTTTACAGTACCT 1283
Qy 1737 GCGTTATACGAGACACCTTATACCTGATACCCCTAACGACAAATAA 1782
Db 1284 GCGTTATACGAGACACCTTATACCTGATACCCCTAACGACAAATAA 1329

113194
LOCUS 113194 1335 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 14 from patent US 5434073.
ACCESSION 113194
VERSION 113194.1 GI:910542
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1335)
AUTHORS Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 14 18-JUL-1995;
FEATURES Location/Qualifiers
1..1335
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 68.5%; Score 1220.4; DB 6; Length 1335;
Best Local Similarity 98.7%; Pred. No. 6.8e-308;
Matches 1230; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 537 TATTGCTGGACTGAGTGGCTGCTAGACCGTCCATCTGTCAAACAACAGCCAAATCGTTGT 596
Db 84 TATTGCTGGACTGAGTGGCTGCTAGACCGTCCATCTGTCAAACAACAGCCAAATAGTTGT 143
Qy 597 TAGCGTTCTGCTACTGTTGGAGGACGAATCAAGACATTAGTCTTAATTTTTGAAAT 656
Db 144 TAGCGTTCTGCTACTGTTGGAGGACGAATCAAGACATTAGTCTTAATTTTTGAAAT 203
Qy 657 CGATCTAACATCACGACCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAAATC 716
Db 204 TGACCTAACATCACGACCTGCTCATGGAGGAAGACAGACGAGCTTAAAGTCCAAAATC 263
Qy 717 AAAACCAATTTGCTACTGATAGTGGCGCATGTGTACATAAACTTGAAGAAGCTGACTTACT 776
Db 264 AAAACCAATTTGCTACTGATAGTGGCGCATGCCACATAACTTGAAGAAGCTGACTTACT 323
Qy 777 AAAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGT 836
Db 324 AAAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGT 383
Qy 837 CATGTATTTGCAAGCGATGCAACATTTATGTATCGAAACGCGAAGGTCCTACTTTGCTGA 896
Db 384 CATGTATTTGCAAGCGATGCAACATTTATGTATCGAAACGCGAAGGTCCTACTTTGCTGA 443
Qy 897 CAAAGATGTTCCGTAACCTTTCGCGACCCACCTGTCGAGAAATTTTGCTAAGCGGACA 956
Db 444 CAAAGATGTTCCGTAACCTTTCGCGACCCACCTGTCGAGAAATTTTGCTAAGCGGACA 503
Qy 957 TGTGCGGCTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGT 1016
Db 504 TGTGCGGCTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGT 563
Qy 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 1076
Db 564 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGTCTCAA 623
Qy 1077 AGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1136
Db 624 AGATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 683
Qy 1137 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTA 1196
Db 684 AGCTCAAGCAACAAGCAATTTTAAACAAAAACCCATCCAGGCTATACGATTTATGAACGTA 743
Qy 1197 CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCA 1256
Db 744 CTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCA 803
Qy 1257 GTTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCT 1316

Query	Match	Score	1220.4	DB 6	Length	1512
Db	804	GTTTACTTACATGTCGGAACAAAGCTTATGAGATCAATATAAAATCTCGTGCT	863			
Qy	1317	GAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG	1376			
Db	864	GAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG	923			
Qy	1377	GGAAAAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAAACGTGTCAACATCAATACGT	1436			
Db	924	GGAAAAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAAACGTGTCAACATCAATACGT	983			
Qy	1437	TGATGTCGATACCAACCAATGCTTAAAGAGTGCACGCTCTTACACGTAGCGAACGTAA	1496			
Db	984	TGATGTCGATACCAACCAATGCTTAAAGAGTGCACGCTCTTACACGTAGCGAACGTAA	1043			
Qy	1497	CTTAGACTTTACAGAGATTATACGATCCCTCGTGATAAGGCTTAAACCTTACTTACAACTCT	1556			
Db	1044	CTTAGACTTTACAGAGATTATACGATCCCTCGTGATAAGGCTTAACTTCTTACAACTCT	1103			
Qy	1557	CGATGCTTTTGGTATATGAGCTATACCTTAACTTGGAAAAAGTAGAGGATATCAACGATGA	1616			
Db	1104	CGATGCTTTTGGTATATGAGCTATACCTTAACTTGGAAAAAGTAGAGGATATCAACGATGA	1163			
Qy	1617	CACCAACCGGTATCATACCGTTTATATGCGCAAGCGACCCCGAAGGAGAGATGCTAGCTA	1676			
Db	1164	CACCAACCGGTATCATACCGTTTATATGCGCAAGCGACCCCGAAGGAGAGATGCTAGCTA	1223			
Qy	1677	TCATTTAGCCCTATGATAAAGATCGTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACCT	1736			
Db	1224	TCATTTAGCCCTATGATAAAGATCGTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACCT	1283			
Qy	1737	CGGTTATACAGGGACACCTATACCTGATATAACCTTAACGACCAATATA	1782			
Db	1284	CGGTTATACAGGGACACCTATACCTGATATAACCTTAACGACCAATATA	1329			
RESULT	15					
LOCUS	A20016					
DEFINITION	SEQ ID NO: 24; Nucleotide sequence for streptokinase fused to yeast alpha-factor.					
ACCESSION	A20016					
VERSION	A20016.1	GI:1247850				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	other sequences; artificial sequences.					
REFERENCE	1 (bases 1 to 1512)					
AUTHORS						
TITLE	PROTEINS AND NUCLEIC ACIDS					
JOURNAL	Patent: WO 9109125-A 24 27-JUN-1991;					
FEATURES	Location/Qualifiers					
source	1..1512					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	7..1506					
	/codon_start=1					
	/transl_table=11					
	/product="streptokinase fused to yeast alpha-factor"					
	/protein_id="CAA01487.1"					
	/db_xref="GI:1247851"					
	/translation="MRPPSFTAVLFAASSALAAPVNTTTEDETAQIPAEAVTGYLDL					
	EGPFDVAVLPFSNSTNGLLFIINTTASIAAKEGVSLDKRIAGPEWLLDRPSVNNQ					
	LTVSVAGTVEGTQDILSKFEIDLSPRHAGKTEGGLSPKSKPATDSGAMPHKLE					
	KADLLKARQLTANVHNSDDYFEVIDFASDATTIDNRGVYFADKXGVSLLTPTQPV					
	EFLLSGHVRVRYKPEKPINQAKSVDVEYTVQFTPLNDDDFRGLKDTKLKTLAIG					
	DTTTSOELLAAQASILNKTHPGTIYERDSSIYTHNDPIFRLLPMQDEFTYHVKRE					
	QAYEINKKSGNEEINLTDLISEKYVYLKGEKPYDPDRSHLKLPTIKYVDVNTNEL					
	LKSEQLLTASERNLDFDLVDPDKAKLLVNLDAFGIMDYTLTGKVEDNDDTNIIL					
	TYVMGKRPEGENASYHLAYDKDRTHEEREVSYLRYTGTPIDPNPDK"					
ORIGIN						

Best Local Similarity 98.3%; Pred. No. 6.8e-308;				
Matches 1233; Conservative 0; Mismatches 21; Indels 0; Gaps 0;				
QY	529	GATGTTCTGTTATTTGCTTGACCTCTGAGTGGCTGCTAGACCGTGCATCTGTCAACAACAGCAAA	588	
Db	253	GATAAAAGAATTTGCTGGACCTGAGTGGCTGCTAGACCGTGCATCTGTCAACAACAGCAAA	312	
QY	589	TTGGTTCTGTAGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTT	648	
Db	313	TTAGTTGTTAGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTT	372	
QY	649	TTTGAAATTCGATCTAAACATACGACCTGCTCATCGGAGGAAAGACAGACGGCTTAAAGT	708	
Db	373	TTTGAAATTTGACCTTAAACATACGACCTGCTCATCGGAGGAAAGACAGACGGCTTAAAGT	432	
QY	709	CCAAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATTTGAGAAAGCT	768	
Db	433	CCAAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATTTGAGAAAGCT	492	
QY	769	GACTTACTAAAGGCTATTCAAGAACAAATTCGCTAAACGTCACAGTAACGACGACTAC	828	
Db	493	GACTTACTAAAGGCTATTCAAGAACAAATTCGCTAAACGTCACAGTAACGACGACTAC	552	
QY	829	TTTGAGGTCATTGATTTTGAAGGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTAC	888	
Db	553	TTTGAGGTCATTGATTTTGAAGGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTAC	612	
QY	889	TTTGTCTGACAAAGATGCTCGGTAAACCTTCGCCGACCCAACTGTCCTCAAGAAATTTTGCCTA	948	
Db	613	TTTGTCTGACAAAGATGCTCGGTAAACCTTCGCCGACCCAACTGTCCTCAAGAAATTTTGCCTA	672	
QY	949	AGCGGACATGTGCGCGTTAGACCATATAAAGAAAAAACCAATACAAAACCAAGCGAAATCT	1008	
Db	673	AGCGGACATGTGCGCGTTAGACCATATAAAGAAAAAACCAATACAAAACCAAGCGAAATCT	732	
QY	1009	GTTGATGTGGAATATACTGTACAGTTTATCTCCCTTAAACCTGATGACGATTTACAGACCA	1068	
Db	733	GTTGATGTGGAATATACTGTACAGTTTATCTCCCTTAAACCTGATGACGATTTACAGACCA	792	
QY	1069	GGTCTCAAGACTATAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAA	1128	
Db	793	GGTCTCAAGACTATAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAA	852	
QY	1129	GAATTTACTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTAT	1188	
Db	853	GAATTTACTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACGATTTAT	912	
QY	1189	GAAAGTGAATCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATG	1248	
Db	913	GAAAGTGAATCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATG	972	
QY	1249	GATCAAGAGTTTACTTACCGTGTATAAAATCGGGAAACAGCTTATAGGATCAATATAAAA	1308	
Db	973	GATCAAGAGTTTACTTACCGTGTATAAAATCGGGAAACAGCTTATAGGATCAATATAAAA	1032	
QY	1309	CTGTGCTCTGAATGAAGAAAAATAAACACACTGACTGATCTCTGAGAAATATTTACGTCCTT	1368	
Db	1033	CTGTGCTCTGAATGAAGAAAAATAAACACACTGACTGATCTCTGAGAAATATTTACGTCCTT	1092	
QY	1369	AAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTCAACCATC	1428	
Db	1093	AAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTCAACCATC	1152	
QY	1429	AAATACGTTGATGTCGATACCAACGAAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGC	1488	
Db	1153	AAATACGTTGATGTCGATACCAACGAAATTTGCTAAAAAGTGAGCAGCTCTTTAAACAGCTAGC	1212	
QY	1489	GAACGTAACCTTAGACTTTCAGAGATTTATACGATCCTCGTGAATAGGCTAAACTACTCTAC	1548	
Db	1213	GAACGTAACCTTAGACTTTCAGAGATTTATACGATCCTCGTGAATAGGCTAAACTACTCTAC	1272	
QY	1549	AACAATCTCGATGCTTTTGGTATTTATGGACTATACCTTAACTGGAAGTAGAGGATAAT	1608	

Db	1273	AACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAAAGTAGAAGATAAT	1332
Qy	1609	CACGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAT	1668
Db	1333	CACGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGGAGAGAAT	1392
Qy	1669	GCTAGCTATCATTTTAGCCTATGATAAAGATCGTTATACCGAAAGAAACGAGAGTTTAC	1728
Db	1393	GCTAGCTATCATTTTAGCCTATGATAAAGATCGTTATACCGAAAGAAACGAGAGTTTAC	1452
Qy	1729	AGCTACCTGCGTTATACAGGACACCTATACCTGATACCCCTAACGACAAATAA	1782
Db	1453	AGCTACCTGCGTTATACAGGACACCTATACCTGATACCCCTAACGACAAATAA	1506

Search completed: February 1, 2006, 04:17:48
Job time : 9060.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 1077.25 Seconds
(without alignments)
11024.772 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tcgcttcacgttcgctcgcg.....ataacctaagcacaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1782	100.0	1782	3	Aaa37642 Chimeric
2	1684	94.5	2096	3	Aaa37643 Chimeric
3	1245.8	69.9	2385	2	Aax80497 Streptoki
4	1245	69.9	1377	3	Aaa37622 Streptoki
5	1243.4	69.8	1245	3	Aaa37633 S. equisi
6	1241.8	69.7	2030	2	Aaa37633 S. equisi
7	1240.4	69.6	1242	2	Aax80492 Streptoco
8	1240.4	69.6	1254	6	Abao5546 Streptoki
9	1240.4	69.6	8993	6	Abao5547 Maxadil
10	1238.8	69.5	1242	2	Aax16632 Streptoco
11	1235.6	69.3	1242	2	Aax16633 Streptoco
12	1232.2	69.1	1327	3	Aaa37628 Streptoki
13	1230.4	69.0	2566	2	Aat77778 Coding se
14	1223.4	68.7	1458	2	Aaql2162 Factor xa
15	1222.6	68.6	7057	12	Adm01294 Plasmid p
16	1220.4	68.5	1335	2	Aaql2156 Streptoki
17	1220.4	68.5	1512	2	Aaql2158 Streptoki
18	1220.2	68.5	2589	2	Aaql2160 OmpAL str
19	1219.4	68.4	1245	2	Aaq20665 SKC-2 str

20	1216.4	68.3	1467	2	AAQ12490	Aaql2490 Factor Xa
21	1212.4	68.0	1473	2	AAQ05603	Aaql05603 Streptoki
22	1211.4	68.0	1407	1	Aan70106	Aan70106 DNA encod
23	1202.8	67.5	2568	1	AAN50493	Aan50493 Sequence
24	1198	67.2	1323	2	AAT29961	Aat29961 Vector pS
25	1189.6	66.8	1242	5	AAF82144	Aaf82144 Mutant st
26	1184	66.4	1209	3	AAZ99249	Aaz99249 DNA encod
27	1181.4	66.3	1245	3	AAZ99250	Aaz99250 DNA encod
28	1150.2	64.5	1661	3	AAA37637	Aaa37637 Chimeric
29	1134.2	63.6	1541	3	AAA37644	Aaa37644 Chimeric
30	1093.6	61.4	1122	3	AAZ99251	Aaz99251 DNA encod
31	1093.6	61.4	1158	3	AAZ99252	Aaz99252 DNA encod
32	1081.2	60.7	1323	13	ADR83811	Adr83811 S. pyogen
33	1081	60.7	1119	2	AAQ12159	Aaql2159 Truncated
34	1080.8	60.7	2253	2	AAQ12161	Aaql2161 Met-core
35	1079.6	60.6	1473	2	AAQ05604	Aaql05604 Streptoki
36	1078.2	60.5	1320	6	ABN70192	Abn70192 Streptoco
37	1068	59.9	1068	2	AAX80493	Aax80493 Recombina
38	1063.4	59.7	2208	2	AAX83589	Aax83589 Recombina
39	984.6	55.3	1245	10	ADP48644	Adp48644 Streptoco
40	829	46.5	1262	2	AAQ10230	Aaql0230 Synthetic
41	450	25.3	450	2	AAX80494	Aax80494 Recombina
42	233.6	13.1	4727	4	AAD09265	Aad09265 pT7-TACS
43	232.8	13.1	1929	13	ADS17488	Adsl7488 Nucleotid
44	232.8	13.1	1929	13	ADR97657	Adr97657 Human fib
45	232.8	13.1	1929	14	ADM44479	Adw44479 Human fib

ALIGNMENTS

RESULT 1
AAA37642
ID AAA37642 standard; DNA; 1782 BP.

AC AAA37642;
15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX Chimeric SK-FBD coding sequence.

DB Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.

XX EPI024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;

XX WPI; 2000-516032/47.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.

XX Example 5; Fig 21b; 58pp; English.

XX This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid

CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Query Match 100.0%; Score 1782; DB 3; Length 1782;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCGCTTCACGTTCCGTCGCGTATCGGTGATTCATTCGTCTAACAGTAAGGCAACCCCGC	60
Db	1	TCGCTTCACGTTCCGTCGCGTATCGGTGATTCATTCGTCTAACAGTAAGGCAACCCCGC	60
Qy	61	CAGCTACCGGTCCTCAAGACAGGAGCAGATCATCGGCACCGTCGCCAGGACCCA	120
Db	61	CAGCTACCGGTCCTCAAGACAGGAGCAGATCATCGGCACCGTCGCCAGGACCCA	120
Qy	121	ACGCTGCCGAGATCTCGATCCGCGAATTAATACGATCTCACTATAGGAGACCAAC	180
Db	121	ACGCTGCCGAGATCTCGATCCGCGAATTAATACGATCTCACTATAGGAGACCAAC	180
Qy	181	GGTTTCCCTCTAGAAATTAATTTGTTTAACTTTAAGAGGAGATATACCATGGTCAAGC	240
Db	181	GGTTTCCCTCTAGAAATTAATTTGTTTAACTTTAAGAGGAGATATACCATGGTCAAGC	240
Qy	241	ACACAGATCTGACCATAGCTGAGAGTGTGTTGATCATGCTGCGGACCTCTATGT	300
Db	241	ACACAGATCTGACCATAGCTGAGAGTGTGTTGATCATGCTGCGGACCTCTATGT	300
Qy	301	GGTCGGAGAAACGTGGGAGAGCGAGCGGACGCATCACTTGCACTTCTAGAAATAGATGC	360
Db	301	GGTCGGAGAAACGTGGGAGAGCGAGCGGACGCATCACTTGCACTTCTAGAAATAGATGC	360
Qy	361	AACGATCAGGACACAAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATAAT	420
Db	361	AACGATCAGGACACAAAGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATAAT	420
Qy	421	CGAGAAACCTGCTCAGTGATCTGCA CAGGCAACCGCGGAGAGGTGGAATGTGAG	480
Db	421	CGAGAAACCTGCTCAGTGATCTGCA CAGGCAACCGCGGAGAGGTGGAATGTGAG	480
Qy	481	AGGCACACCTCTGTGAGACACATCTGAGCGGATCTGGCCCTTCAACCGATGTTGCTATT	540
Db	481	AGGCACACCTCTGTGAGACACATCTGAGCGGATCTGGCCCTTCAACCGATGTTGCTATT	540
Qy	541	GCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCAACAAACAGCCAAATGGTGTAGC	600
Db	541	GCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCAACAAACAGCCAAATGGTGTAGC	600
Qy	601	GTTCGTGTAATCTGTTGAGGAGAGATCAAGCATTTAGTCTTAAATTTTTTGAATTCGAT	660
Db	601	GTTCGTGTAATCTGTTGAGGAGAGATCAAGCATTTAGTCTTAAATTTTTTGAATTCGAT	660
Qy	661	CTAACATCACACCTGCTCATGGAGGAAGACAGAGCAGGCTTAAGTCCAAATCAAAA	720
Db	661	CTAACATCACACCTGCTCATGGAGGAAGACAGAGCAGGCTTAAGTCCAAATCAAAA	720

Db	661	CTAACATCACACCTGCTCATGGAGGAAGACAGAGCAGGCTTAAGTCCAAATCAAAA	720
Qy	721	CCATTTGCTACTGATAGTGGCGGATGTCATATAAATCTTGAGAAAGCTGACTTACTAAG	780
Db	721	CCATTTGCTACTGATAGTGGCGGATGTCATATAAATCTTGAGAAAGCTGACTTACTAAG	780
Qy	781	GCTATTCAAGAAACAATTTGATCGGTAACTCCACAGTAACGACGACTACTTTGAGGTCAAT	840
Db	781	GCTATTCAAGAAACAATTTGATCGGTAACTCCACAGTAACGACGACTACTTTGAGGTCAAT	840
Qy	841	GATTTTGCAGCGATGCAACATTTACTGATCGAAACGCAAGGTCTACTTTGCTGACAAA	900
Db	841	GATTTTGCAGCGATGCAACATTTACTGATCGAAACGCAAGGTCTACTTTGCTGACAAA	900
Qy	901	GATGGTTCGTTAAGTTCGCGACCCCACTGTCGACAGAAATTTTGTCTAAGGCGCATGTG	960
Db	901	GATGGTTCGTTAAGTTCGCGACCCCACTGTCGACAGAAATTTTGTCTAAGGCGCATGTG	960
Qy	961	CGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTGAA	1020
Db	961	CGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTGAA	1020
Qy	1021	TATACGTGACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAGAT	1080
Db	1021	TATACGTGACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAGAT	1080
Qy	1081	ACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTAGCT	1140
Db	1081	ACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTAGCT	1140
Qy	1141	CAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTTATACGATTTATGACGACTCTCC	1200
Db	1141	CAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTTATACGATTTATGACGACTCTCC	1200
Qy	1201	TCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTT	1260
Db	1201	TCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTT	1260
Qy	1261	ACTTACCGTGTAAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTGAAT	1320
Db	1261	ACTTACCGTGTAAAAAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTGAAT	1320
Qy	1321	GAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTCAGTCTTTAAAAAGGGGAA	1380
Db	1321	GAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTCAGTCTTTAAAAAGGGGAA	1380
Qy	1381	AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTTGAT	1440
Db	1381	AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTTGAT	1440
Qy	1441	GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAATTA	1500
Db	1441	GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAATTA	1500
Qy	1501	GACTTCAGAGATTTATACGATCTCTGATGTAAGGCTTAACTCTCTACCAATCTCGAT	1560
Db	1501	GACTTCAGAGATTTATACGATCTCTGATGTAAGGCTTAACTCTCTACCAATCTCGAT	1560
Qy	1561	GCTTTTGGTATTATGACTATACCTTAACTGAAAAAGTAGAGGATATACGATGACACC	1620
Db	1561	GCTTTTGGTATTATGACTATACCTTAACTGAAAAAGTAGAGGATATACGATGACACC	1620
Qy	1621	AACCGTATCAATACCGTTTATATGGCGAAGCGACCCGAGAGAGGAATGTAGCTATCAT	1680
Db	1621	AACCGTATCAATACCGTTTATATGGCGAAGCGACCCGAGAGAGGAATGTAGCTATCAT	1680
Qy	1681	TTAGCTATCATATAAGATCGTTATACCGAAGAGACGAGAGGTTTACAGTACCTCGGT	1740
Db	1681	TTAGCTATCATATAAGATCGTTATACCGAAGAGACGAGAGGTTTACAGTACCTCGGT	1740
Qy	1741	TATACAGGAGACCTTATACCTGATTAACCTTAAACGACAAATAA	1782
Db	1741	TATACAGGAGACCTTATACCTGATTAACCTTAAACGACAAATAA	1782

RESULT 2
 AAA37643
 ID AAA37643 standard; DNA; 2096 BP.
 XX
 AC AAA37643;
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 XX Chimeric SK-FBD coding sequence.
 XX
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Homo sapiens.
 OS Chimeric.
 XX
 PN EP1024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 XX 23-DEC-1999; 99EP-00310541.
 XX
 XX 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RES.
 XX
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 PI
 DR WPI; 2000-516032/47.
 XX
 PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 PS Example 6; Fig 22b; 59pp; English.
 XX
 CC This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;
 Query Match 94.5%; Score 1684; DB 3; Length 2096;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 1 TCGCTTCACGTTTCGGTTCGGTATCGGTGATTCATTCTGCTAACCAAGTAAAGCAACCCCGC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 51 TCGCTTCACGTTTCGGTTCGGTATCGGTGATTCATTCTGCTAACCAAGTAAAGCAACCCCGC 110
 Qy 61 CAGCTAGCCGGTCTCTCAACGACAGGAGCAGATCATGCGCACCCGTCGCGCAGCAACCA 120
 Db 111 CAGCTAGCCGGTCTCTCAACGACAGGAGCAGATCATGCGCACCCGTCGCGCAGCAACCA 170
 Qy 121 ACGTGCCTCCTAGAAATAATTTTGTAACTTTAAGAGGAGATATACCATGGTGAAGC 240
 Db 171 ACGTGCCTCCTAGAAATAATTTTGTAACTTTAAGAGGAGATATACCATGGTGAAGC 290
 Qy 231 GGTTCCTCTAGAAATAATTTTGTAACTTTAAGAGGAGATATACCATGGTGAAGC 300
 Db 241 ACAAAGATTTGACCATAGCTGAGAAGTGTTCATCATGCTGCGGACCTTCCTATGT 350
 Qy 291 ACAAAGATTTGACCATAGCTGAGAAGTGTTCATCATGCTGCGGACCTTCCTATGT 360
 Db 301 GGTTCGAGAAACGTGGGAGAGGCGGACGCGATCATCTTGCATCTTAGAAATAGATGC 410
 Qy 351 GGTTCGAGAAACGTGGGAGAGGCGGACGCGATCATCTTGCATCTTAGAAATAGATGC 420
 Db 361 AACGATCAGGACACAGGACATCTTATAGATTGAGACACCTGGAGCAAGAGATAT 470
 Qy 411 AACGATCAGGACACAGGACATCTTATAGATTGAGACACCTGGAGCAAGAGATAT 480
 Db 421 CGAGAAACCTGCTCCAGTGCATCTGCACAGCAAGCGGAGGAGAGTGGAGTGTGAG 530
 Qy 471 CGAGAAACCTGCTCCAGTGCATCTGCACAGCAAGCGGAGGAGAGTGGAGTGTGAG 540
 Db 481 AGGCACACCTGCTGCAGACACCATCTGCGCGGATCTGCGCCCTTCAACGATGTCGTATT 590
 Qy 531 AGGCACACCTGCTGCAGACACCATCTGCGCGGATCTGCGCCCTTCAACGATGTCGTATT 600
 Db 541 GCTGCACCTGAGTGGTCTAGACCGTCCATCTGTCAACACAGCAAAATGGTGTAGC 650
 Qy 591 GCTGCACCTGAGTGGTCTAGACCGTCCATCTGTCAACACAGCAAAATGGTGTAGC 660
 Db 601 GTTCTGTGTACTGTGTGAGGAGCAAGATCAAGACATTAGTCTTAATAATTTTGAATCGAT 710
 Qy 651 GTTCTGTGTACTGTGTGAGGAGCAAGATCAAGACATTAGTCTTAATAATTTTGAATCGAT 720
 Db 661 CTAACATCAGCACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCAAAAATCAAAA 770
 Qy 711 CTAACATCAGCACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCAAAAATCAAAA 780
 Db 721 CCAATTTGCTACTGATAGTGGCGCATGTGCATATAAACTTTGAGAAAGCTGACTTAAAG 830
 Qy 771 CCAATTTGCTACTGATAGTGGCGCATGTGCATATAAACTTTGAGAAAGCTGACTTAAAG 840
 Db 781 GCTATTCAGAAACAAATTGATCGCTAAACGTCCACAGTAACGACGACTATTTGAGGTCAAT 890
 Qy 831 GCTATTCAGAAACAAATTGATCGCTAAACGTCCACAGTAACGACGACTATTTGAGGTCAAT 900
 Db 841 GATTTTCAAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGACAAA 950
 Qy 891 GATTTTCAAGCGATGCAACCATTTACTGATCGAAACGCAAGGCTTACTTTGCTGACAAA 960
 Db 901 GATGGTTTCGGTAAACCTTGGCCGACCAACCTGCTCCAGAAATTTTGTAAAGCGGACATGTG 1010
 Qy 951 GATGGTTTCGGTAAACCTTGGCCGACCAACCTGCTCCAGAAATTTTGTAAAGCGGACATGTG 1020
 Db 961 CGGTTTACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTGAA 1070
 Qy 1011 CGGTTTACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTGAA 1080
 Db 1021 TATGCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAAGGCTTCAAGAT 1130
 Qy 1071 TATGCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAAGGCTTCAAGAT 1140
 Db 1081 ACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTAGCT 1190
 Qy 1131 ACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTAGCT

QY 1141 CAAGCACAAAGCATTTTAAACAAAACACCCAGGCTATACGATTTATGAACGCTACTCC 1200
 DB |||||
 QY 1191 CAAGCACAAAGCATTTTAAACAAAACACCCAGGCTATACGATTTATGAACGCTACTCC 1250
 DB |||||
 QY 1201 TCAATCGTCATCATGACATGACATTTTCCGTCAGTATTTTACCAATGGAATCAAGATTTT 1260
 DB |||||
 QY 1251 TCAATCGTCATCATGACATGACATTTTCCGTCAGTATTTTACCAATGGAATCAAGATTTT 1310
 DB |||||
 QY 1261 ACTTACCGTGTAAATAACGGAACAGCTTATAGGATCAATAAAAATCTGCTCGAAT 1320
 DB |||||
 QY 1311 ACTTACCGTGTAAATAACGGAACAGCTTATAGGATCAATAAAAATCTGCTCGAAT 1370
 DB |||||
 QY 1321 GAAGAAATAAACCAACACAGCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGGGAA 1380
 DB |||||
 QY 1371 GAAGAAATAAACCAACAGCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGGGAA 1430
 DB |||||
 QY 1381 AAGCGTATGATCCCTTTGATCGAGTCACTTTGAAACTGTTCAACCATCAAAATACGTTGAT 1440
 DB |||||
 QY 1441 GTCGATACCAACGAATGCTTAAAAAGTGAGAGCTCTTAAACAGCTAGCGAAACGTAACCTTA 1500
 DB |||||
 QY 1491 GTCGATACCAACGAATGCTTAAAAAGTGAGAGCTCTTAAACAGCTAGCGAAACGTAACCTTA 1550
 DB |||||
 QY 1501 GACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAAACATCTCGAT 1560
 DB |||||
 QY 1551 GACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAAACATCTCGAT 1610
 DB |||||
 QY 1561 GCTTTTGGTATTATGAGCTATACCTTAACTGGAAGATAGAGGATTAATCACGATGACACC 1620
 DB |||||
 QY 1611 GCTTTTGGTATTATGAGCTATACCTTAACTGGAAGATAGAGGATTAATCACGATGACACC 1670
 DB |||||
 QY 1621 AACCGTATCAATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTATCAT 1680
 DB |||||
 QY 1671 AACCGTATCAATACCGTTTATATGCGCAAGCGACCCGAAAGGAGAGAAATGCTAGCTATCAT 1730
 DB |||||
 QY 1681 TTAGCTATGAT 1692
 DB |||||
 QY 1731 TTAGCTGGTGGT 1742
 DB |||||

RESULT 3

AXX80497
 ID AAX80497 standard; cDNA; 2385 BP.
 XX AC
 AC AAX80497;
 XX DT
 DT 26-AUG-1999 (first entry)
 XX DE
 DE Streptokinase and maltose binding protein fusion protein encoding cDNA.
 XX KW
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW nSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX OS
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX OS
 OS Synthetic.
 XX PN
 PN W09931247-A1.
 XX XX
 XX 24-JUN-1999.
 XX XX
 XX 15-DEC-1998; 98WO-US026694.
 XX PF
 PF 15-DEC-1997; 97US-0069497P.
 XX PR
 PR (HARD) HARVARD COLLEGE.
 XX FA
 FA Reed GL;
 XX PI
 PI WPI; 1999-395183/33.
 XX DR

DR P-PSDB; AAY24797.
 XX N-terminally deleted streptokinase.
 PT
 XX
 PS Example; Page 45-48; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC encodes a streptokinase and maltose binding protein fusion protein from
 CC an example of the present invention
 XX
 SQ Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;
 Query Match 69.9%; Score 1245.8; DB 2; Length 2385;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1250; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 523 TTCAACCGATGTTTCGATTTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAC 582
 DB |||||
 QY 583 AGCCAAATGGTTTGTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTTAGTCTT 642
 DB |||||
 QY 1189 AGCCAAATTTAGTTTGTAGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTTAGTCTT 1248
 DB |||||
 QY 643 AAATTTTGGAAATCGATCTTAACATCACGACCTGCTCATCGAGGAAGACAGACGAGGC 702
 DB |||||
 QY 1249 AAATTTTGGAAATCGATCTTAACATCACGACCTGCTCATCGAGGAAGACAGACGAGGC 1308
 DB |||||
 QY 703 TTAAGTCCAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 762
 DB |||||
 QY 1309 TTAAGTCCAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1368
 DB |||||
 QY 763 AAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAACGAC 822
 DB |||||
 QY 1369 AAAGCTGACTTACTAAAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAACGAC 1428
 DB |||||
 QY 823 GACTACTTTGAGGTCAATTTGATTTTGCAGCGATGCAACATTTACTGTATCGAAACGCGAAG 882
 DB |||||
 QY 1429 GACTACTTTGAGGTCAATTTGATTTTGCAGCGATGCAACATTTACTGTATCGAAACGCGAAG 1488
 DB |||||
 QY 883 GTCTACTTTGCTGACAAAGATGTTGCTGTAACCTTCCGCGACCCCAACCTGTCCTCAAGAAATTT 942
 DB |||||
 QY 1489 GTCTACTTTGCTGACAAAGATGTTGCTGTAACCTTCCGCGACCCCAACCTGTCCTCAAGAAATTT 1548
 DB |||||
 QY 943 TTGCTAAGCGGACATGTGCGCGTTTAGACCATATAAAGAAAAAACAATACAAAACCAAGCG 1002
 DB |||||
 QY 1549 TTGCTAAGCGGACATGTGCGCGTTTAGACCATATAAAGAAAAAACAATACAAAACCAAGCG 1608
 DB |||||
 QY 1003 AAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGCATTTTC 1062
 DB |||||
 QY 1609 AAATCTGTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGCATTTTC 1668
 DB |||||
 QY 1063 AGACCAAGGTCTCAAGACATCTAAGCTATTGAAACACACTAGCTATCGGTGACCATCACA 1122
 DB |||||

Db 1669 AGACAGGCTCTCAAGATACCTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCACA 1728
Qy 1123 TCTCAAGAAATTACTAGCTCAAGCACAAGCAATTTTAAACAAAACACCCAGGCTATACG 1182
Db 1729 TCTCAAGAAATTACTAGCTCAAGCACAAGCAATTTTAAACAAAACACCCAGGCTATACG 1788
Qy 1183 ATTTATGAACGTGACTCTCAATCGTCACTCATGACATGACATGCAATTTTCCGTACGATTTTA 1242
Db 1789 ATTTATGAACGTGACTCTCAATCGTCACTCATGACATGACATGCAATTTTCCGTACGATTTTA 1848
Qy 1243 CCAATGGATCAAGAGTTTACTTACCGTGTGTTAAATAATCGGAAACAGCTTATAGATCAAT 1302
Db 1849 CCAATGGATCAAGAGTTTACTTACCGTGTGTTAAATAATCGGAAACAGCTTATAGATCAAT 1908
Qy 1303 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1362
Db 1909 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1968
Qy 1363 GTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAACTGTTTC 1422
Db 1969 GTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAACTGTTTC 2028
Qy 1423 ACCATCAATAGCTGATCTGATACCAACGATTCCTTAAAGTGAAGCTCTTTAACA 1482
Db 2029 ACCATCAATAGCTGATCTGATACCAACGATTCCTTAAAGTGAAGCTCTTTAACA 2088
Qy 1483 GCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTA 1542
Db 2089 GCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTA 2148
Qy 1543 CTCTCAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 1602
Db 2149 CTCTCAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 2208
Qy 1603 GATATACGATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGA 1662
Db 2209 GATATACGATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGA 2268
Qy 1663 GAGATGCTAGCTATCATTTAGCCTATGATTAAGATCGTTATACCGAAGAACGAGAA 1722
Db 2269 GAGATGCTAGCTATCATTTAGCCTATGATTAAGATCGTTATACCGAAGAACGAGAA 2328
Qy 1723 GTTTACAGTACTCGTGGTTATACAGGACACCTTATACCTGATTAACCTTACGACAAA 1779
Db 2329 GTTTACAGTACTCGTGGTTATACAGGACACCTTATACCTGATTAACCTTACGACAAA 2385

RESULT 4
AAA37622
ID AAA37622 standard; DNA; 1377 BP.
XX AC AAA37622;
XX 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX DE Streptokinase-NTRN gene.
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
XX plasminogen; human; fibrinectin; thrombolytic therapy;
XX cardiovascular disorder; ss.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS EP1024192-A2.
XX PN 02-AUG-2000.
XX PD 23-DEC-1999; 99EP-00310541.
XX PF 24-DEC-1998; 98IN-DE003825.
XX PR (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX WPI; 2000-516032/47.
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX Example 1; Fig 11; 58pp; English.
XX This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN
CC stands for N-terminally repaired with native sequence). The invention
CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
CC activation, and fibrin binding regions of human fibrinectin, which are
CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
CC possesses the ability to bind with fibrin independently and also
CC characteristically retains a PG activation ability which becomes evident
CC only after a pronounced duration, or lag, after exposure of the PA to a
CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
CC domain polypeptides are useful in thrombolytic therapy for various kinds
CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
CC as well as kinetics of plasminogen activation that are distinct from that
CC of natural streptokinase in being characterised by a temporary delay, or
CC lag of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX SQ Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;

Query Match 69.9%; Score 1245; DB 3; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 538 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATGTTGTTT 597
Db 133 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATGTTGTTT 192
Qy 598 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 657
Db 193 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 252
Qy 658 GATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 717
Db 253 GATCTAATCATCAGACCTGCTCATGGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 312
Qy 718 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAGCTGACTACTA 777
Db 313 AAACCATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAGCTGACTACTA 372
Qy 778 AAGGCTATTCAAGAACAAATGATGCTTAACGTCCAAGTACACGACGACTACTTTGAGTC 837
Db 373 AAGGCTATTCAAGAACAAATGATGCTTAACGTCCAAGTACACGACGACTACTTTGAGTC 432
Qy 838 ATTGATTTTGCAGCGATGCAACCATTAATGATCGAAGCAAGGCTTACTTTGCTGAC 897
Db 433 ATTGATTTTGCAGCGATGCAACCATTAATGATCGAAGCAAGGCTTACTTTGCTGAC 492
Qy 898 AAGATGTTTGGTAACTTTGCCGACCCCACTCTGCCAAGAAATTTTTCGTAAGCGGACAT 957
Db 493 AAGATGTTTGGTAACTTTGCCGACCCCACTCTGCCAAGAAATTTTTCGTAAGCGGACAT 552
Qy 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 553 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 612

QY 1018 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 1077
 Db 613 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 672
 QY 1078 GATACCTAGCTATTGAACACCTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 1137
 Db 673 GATACCTAGCTATTGAACACCTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 732
 QY 1138 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 1197
 Db 733 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 792
 QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 1257
 Db 793 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 852
 QY 1258 TTTACTTACCGTGTAAATATCGGGAACAGCTTTATAGATCAATAAAAATCTGCTCTG 1317
 Db 853 TTTACTTACCGTGTAAATATCGGGAACAGCTTTATAGATCAATAAAAATCTGCTCTG 912
 QY 1318 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACCTCTTAAAAAGGG 1377
 Db 913 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACCTCTTAAAAAGGG 972
 QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAATACGTT 1437
 Db 973 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAATACGTT 1032
 QY 1438 GATGTCGATACCAAGAAATGCTTAAAGTAGAGCTCTTAAACAGCTAGCGGAACCTAAC 1497
 Db 1033 GATGTCGATACCAAGAAATGCTTAAAGTAGAGCTCTTAAACAGCTAGCGGAACCTAAC 1092
 QY 1498 TTAGACTTCAGAGATTATACGATCTCTGATAGGCTTAACTACTCTACACATCTC 1557
 Db 1093 TTAGACTTCAGAGATTATACGATCTCTGATAGGCTTAACTACTCTACACATCTC 1152
 QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC 1617
 Db 1153 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATATACAGATGAC 1212
 QY 1618 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
 Db 1213 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1272
 QY 1678 CATTTAGCTTATGATAAGATCGTTATACGGAAGACGAGAGTTTACAGTACCTG 1737
 Db 1273 CATTTAGCTTATGATAAGATCGTTATACGGAAGACGAGAGTTTACAGTACCTG 1332
 QY 1738 CGTTATACAGGACACCTTACTCTGATACCCCTTAACGACAAATAA 1782
 Db 1333 CGTTATACAGGACACCTTACTCTGATACCCCTTAACGACAAATAA 1377

RESULT 5
 AAA37633 standard; DNA; 1245 BP.

AC AAA37633;
 XX
 AC
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)

DE S. equisimilis streptokinase coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers
 FH 1..1245
 CDS /*tag= a

/product= "streptokinase"

FT
 XX EP1024192-A2.
 XX
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 XX
 XX 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX
 DR WPI; 2000-516032/47.
 DR P-PSDB; AAY90282.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.

XX Example 3; Fig 3; 58pp; English.

XX This sequence represents the human Streptococcus equisimilis
 CC streptokinase coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match 69.8%; Score 1243.4; DB 3; Length 1245;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACACGCCAATGTTGTT 597
 Db 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACACGCCAATGTTGTT 60
 QY 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 657
 Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 120
 QY 658 GATCTAACATCAGCAGCTGCTCATGGAGGAAAGACAGAGCAGGCTTAAGTCCAAATCA 717
 Db 121 GATCTAACATCAGCAGCTGCTCATGGAGGAAAGACAGAGCAGGCTTAAGTCCAAATCA 180
 QY 718 AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 777
 Db 181 AAACCAATTTGCTACTCATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 240
 QY 778 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGGTC 837
 Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGGTC 300

QY 838 ATTGATTTTGCAGCGATGCAACCATTAATGATCGAAACGCAAGGCTACTTTGCTGAC 897
 Db |||||
 QY 301 ATTGATTTTGCAGCGATGCAACCATTAATGATCGAAACGCAAGGCTACTTTGCTGAC 360
 Db |||||
 QY 898 AAGATGTTTGGTAACTTGGCGACCAACCTGTCGAAGTATTTTGGCTAAGCGGACAT 957
 Db |||||
 QY 361 AAGATGTTTGGTAACTTGGCGACCAACCTGTCGAAGTATTTTGGCTAAGCGGACAT 420
 Db |||||
 QY 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
 Db |||||
 QY 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
 Db |||||
 QY 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAAGGCTCTCAAA 1077
 Db |||||
 QY 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAAGGCTCTCAAA 540
 Db |||||
 QY 1078 GATACATAAGCTATTGAACACACTAGCTATCGGTGACCAATCATCATCTCAAGAAATTAATA 1137
 Db |||||
 QY 541 GATACATAAGCTATTGAACACACTAGCTATCGGTGACCAATCATCATCTCAAGAAATTAATA 600
 Db |||||
 QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1197
 Db |||||
 QY 601 GCTCAAGCAAAAGCAATTTTAAACAAAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660
 Db |||||
 QY 1198 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
 Db |||||
 QY 661 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
 Db |||||
 QY 1258 TTTACTTACCGTGTAAACCTGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Db |||||
 QY 721 TTTACTTACCGTGTAAACCTGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
 Db |||||
 QY 1318 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
 Db |||||
 QY 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 Db |||||
 QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCAATCAATACGTT 1437
 Db |||||
 QY 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCAATCAATACGTT 900
 Db |||||
 QY 1438 GATGCGATACCAAGAAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC 1497
 Db |||||
 QY 901 GATGCGATACCAAGAAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC 960
 Db |||||
 QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTCTCTCAACAATCTC 1557
 Db |||||
 QY 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGCTTAACTCTCTCAACAATCTC 1020
 Db |||||
 QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 1617
 Db |||||
 QY 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 1080
 Db |||||
 QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
 Db |||||
 QY 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
 Db |||||
 QY 1678 CATTTAGCCTATGATAAGATCGTTTATACGGAAGAGAACGAGAAGTTTACAGCTACCTG 1737
 Db |||||
 QY 1141 CATTTAGCCTATGATAAGATCGTTTATACGGAAGAGAACGAGAAGTTTACAGCTACCTG 1200
 Db |||||
 QY 1738 CGTTATACAGGACACCTTATACCTGATACCCCTTAAACGACAAATAA 1782
 Db |||||
 QY 1201 CGTTATACAGGACACCTTATACCTGATACCCCTTAAACGACAAATAA 1245
 Db |||||

RESULT 6

AAQ11651

ID AAQ11651 standard; DNA; 2030 BP.

XX

AC AAQ11651;

XX

DT 08-JUL-1991 (first entry)

XX

DE FB-FB-SK fusion construct.
 XX Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
 KW streptokinase; fusion protein; ss.
 XX Staphylococcus aureus.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 10..184
 FT /*tag= a
 FT /label= FB monomer
 FT CDS 185..358
 FT /*tag= b
 FT /label= FB monomer
 FT CDS 359..1601
 FT /*tag= c
 FT /label= streptokinase
 XX
 XX US0111686-A.
 XX 30-APR-1991.
 XX 15-NOV-1989; 89US-00437769.
 XX 21-SEP-1987; 87US-00099242.
 XX (CREA-) CREATIVE BIOMOLEC.
 XX Pang RHL;
 XX WPI; 1991-140198/19.
 XX P-PSDB; AAR11829.
 XX
 XX Imparting injectable fibrinolytic agent - with affinity for intravascular
 PT thrombus, by linking agent to fibrin binding domain.
 XX
 XX Disclosure; Fig 5; 18pp; English.

CC The DNA encodes an FB-PB dimer linked to the streptokinase coding
 CC sequence. The FB fragment has selective affinity for fibrin, low affinity
 CC for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
 CC capability. See also AAQ11649 and AAQ11650
 XX
 SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;

Query Match 69.7%; Score 1241.8; DB 2; Length 2030;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGCTCATCTGTCAACAACAGCCAAATGGTGT 597
 Db |||||
 QY 358 ATTGCTGGACCTGAGTGGCTGTAGACCGCTCATCTGTCAACAACAGCCAAATGGTGT 417
 Db |||||
 QY 598 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACATTAGTCTTAAATTTTGAATC 657
 Db |||||
 QY 418 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACATTAGTCTTAAATTTTGAATC 477
 Db |||||
 QY 658 GATCTAATCATCAGCACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATCA 717
 Db |||||
 QY 478 GATCTAATCATCAGCACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATCA 537
 Db |||||
 QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAGCTGACTTACTA 777
 Db |||||
 QY 538 AAACCAATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAGCTGACTTACTA 597
 Db |||||
 QY 778 AAGGCTATTCAAGAACCAATTTGATCGCTTAACTGCTCAAGTACAGGACTTGTGAGTGC 837
 Db |||||
 QY 598 AAGGCTATTCAAGAACCAATTTGATCGCTTAACTGCTCAAGTACAGGACTTGTGAGTGC 657
 Db |||||
 QY 838 ATTGATTTTGCAGGAGGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 897
 Db |||||
 QY 658 ATTGATTTTGCAGGAGGATGCAACCATTTACTGATCGAAACGCAAGGCTCTACTTTGCTGAC 717
 Db |||||

QY 898 AAAGATGGTTCGGTAAACCTTGGCGACCCAACTCTGTCACAGAAATTTTGTCTAAGCGGACAT 957
 Db 718 AAAGATGGTTCGGTAAACCTTGGCGACCCAACTCTGTCACAGAAATTTTGTCTAAGCGGACAT 777
 QY 958 GTGCGGTTAGACCATATTAAGAAAAACCAATACAAAACCAAGCGAAATCTGTGATGTG 1017
 Db 778 GTGCGGTTAGACCATATTAAGAAAAACCAATACAAAACCAAGCGAAATCTGTGATGTG 837
 QY 1018 GAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 1077
 Db 838 GAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAAA 897
 QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
 Db 898 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 957
 QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGCCAGGCTATACGATTTATGAACGTGAC 1197
 Db 958 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGCCAGGCTATACGATTTATGAACGTGAC 1017
 QY 1198 TCCCTCAATCGTCACTCATGACATGACATTTTCCGTGACGATTTTACCATGGATCAAGAG 1257
 Db 1018 TCCCTCAATCGTCACTCATGACATGACATTTTCCGTGACGATTTTACCATGGATCAAGAG 1077
 QY 1258 TTTACTACCGTGTGTTAAATAACCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Db 1078 TTTACTACCGTGTGTTAAATAACCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1137
 QY 1318 AATGAAGAAATTAACAACTGACCTGATCTCTGAGAAATATTAATCGTCTTAAAAAGGG 1377
 Db 1138 AATGAAGAAATTAACAACTGACCTGATCTCTGAGAAATATTAATCGTCTTAAAAAGGG 1197
 QY 1378 GAAGACCGTATGATCCCTTTCATGTCAGCTGACCTTGAACCTGTTCCACCAATAGCTT 1437
 Db 1198 GAAGACCGTATGATCCCTTTCATGTCAGCTGACCTTGAACCTGTTCCACCAATAGCTT 1257
 QY 1438 GATGTCGATACCAACAAATGCTTAAAGTGACGACCTCTTAACAGCTAGCGAACGTAAC 1497
 Db 1258 GATGTCGATACCAACAAATGCTTAAAGTGACGACCTCTTAACAGCTAGCGAACGTAAC 1317
 QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAACTCTCTCAACAATCTC 1557
 Db 1318 TTAGACTTCAGAGATTTATACGATCTCTGATAGGCTTAACTCTCTCAACAATCTC 1377
 QY 1558 GATGCTTTTGGTATTTATGACACTATCTTAACTGGAAGAGTAGAGATATACGATGAC 1617
 Db 1378 GATGCTTTTGGTATTTATGACACTATCTTAACTGGAAGAGTAGAGATATACGATGAC 1437
 QY 1618 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
 Db 1438 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1497
 QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAGACGAGAAGTTTACAGCTACCTG 1737
 Db 1498 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAGACGAGAAGTTTACAGCTACCTG 1557
 QY 1738 CGTTATACGGGACACTATACCTGATTAACCCCTTAACGACAAATPAA 1782
 Db 1558 CGTTATACGGGACACTATACCTGATTAACCCCTTAACGACAAATPAA 1602

RESULT 7

AXX80492
 ID AAX80492 standard; cDNA; 1242 BP.
 XX AC AAX80492;
 XX AC AAX80492;
 DT 17-OCT-2003 (revised)
 DT 26-AUG-1999 (first entry)
 XX
 DE Streptococcus equisimilis native streptokinase encoding cDNA.
 XX
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;

KW nSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX WO9931247-A1.
 XX 24-JUN-1999.
 XX 15-DEC-1998; 98WO-US026694.
 XX 15-DEC-1997; 97US-0069497P.
 XX (HARD) HARVARD COLLEGE.
 XX Reed GL;
 XX WPI; 1999-395183/33.
 DR P-PSDB; AAY24794.
 XX N-terminally deleted streptokinase.
 XX Claim 44; Page 58-60; 73pp; English.
 XX The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (I) encoding a modified bacterial streptokinase; (3) an expression vector comprising (I); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence encodes native streptokinase (nSK). (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;
 Query Match 69.6%; Score 1240.4; DB 2; Length 1242;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCAATTAGTCTTTAAATCTTGAATC 657
 Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCAATTAGTCTTTAAATCTTGAATC 60
 QY 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATCTTGAATC 657
 Db 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTTAAATCTTGAATC 120
 QY 658 GATCTAACATCATCGACCTGCTCATGAGGAGAAAGACAGACGACGCTTAAGTCTTCAAAATCA 717
 Db 121 GATCTAACATCATCGACCTGCTCATGAGGAGAAAGACAGACGACGCTTAAGTCTTCAAAATCA 180
 QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGGAAGCTGACTTACTA 777
 Db 181 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAACTTGAGGAAGCTGACTTACTA 240
 QY 778 AAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 837
 Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 300

QY 838 ATTGATTTTCCAAAGCGATGCAACATTTACTGATCGAAACGGCAAGGTCCTACTTTGCTGAC 897
 Db |||||
 QY 301 ATTGATTTTCCAAAGCGATGCAACATTTACTGATCGAAACGGCAAGGTCCTACTTTGCTGAC 360
 Db |||||
 QY 898 AAGATGGTTTCGGTAAACCTTTCGCGACCCCAACCTCTCAAGAAATTTTTCGTAAGCGGACAT 957
 Db |||||
 QY 361 AAGATGGTTTCGGTAAACCTTTCGCGACCCCAACCTCTCAAGAAATTTTTCGTAAGCGGACAT 420
 Db |||||
 QY 958 GTCGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
 Db |||||
 QY 421 GTCGCGGTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
 Db |||||
 QY 1018 GAATATACTGTAAGTTTACTCCCTTAAACCCCTGATGACGANTTTAGACCAAGGTCCTCAA 1077
 Db |||||
 QY 481 GAATATACTGTAAGTTTACTCCCTTAAACCCCTGATGACGANTTTAGACCAAGGTCCTCAA 540
 Db |||||
 QY 1078 GATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
 Db |||||
 QY 541 GATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
 Db |||||
 QY 1138 GCTCAAGCACAAGCATTTTAAACAAAACCCACCGAGCTATAGGATTTATGAACGTGAC 1197
 Db |||||
 QY 601 GCTCAAGCACAAGCATTTTAAACAAAACCCACCGAGCTATAGGATTTATGAACGTGAC 660
 Db |||||
 QY 1198 TCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
 Db |||||
 QY 661 TCTCAATCGTCACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
 Db |||||
 QY 1258 TTTACTACCGTGTAAATTCGGGAACAAGCTTATAGGATCAATAAAATCTGGTCTG 1317
 Db |||||
 QY 721 TTTACTACCGTGTAAATTCGGGAACAAGCTTATAGGATCAATAAAATCTGGTCTG 780
 Db |||||
 QY 1318 AATGAAGAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
 Db |||||
 QY 781 AATGAAGAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 Db |||||
 QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCATCAAAATACGTT 1437
 Db |||||
 QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCATCAAAATACGTT 900
 Db |||||
 QY 1438 GATGTCGATACCAACGAATTTGTAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
 Db |||||
 QY 901 GATGTCGATACCAACGAATTTGTAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
 Db |||||
 QY 1498 TTAGACTTCAGAGATTTATAGATCTCTCGTGATAGGCTAAACTTACTTACAACTCTC 1557
 Db |||||
 QY 961 TTAGACTTCAGAGATTTATAGATCTCTCGTGATAGGCTAAACTTACTTACAACTCTC 1020
 Db |||||
 QY 1558 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATTAATCAGATGAC 1617
 Db |||||
 QY 1021 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATTAATCAGATGAC 1080
 Db |||||
 QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAGATGCTAGCTAT 1677
 Db |||||
 QY 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1140
 Db |||||
 QY 1678 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAAGTTTACAGTACCTG 1737
 Db |||||
 QY 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAAGTTTACAGTACCTG 1200
 Db |||||
 QY 1738 CGTTATACAGGGACCTTATACCTGATGATACCCCTAACGACAAA 1779
 Db |||||
 QY 1201 CGTTATACAGGGACCTTATACCTGATGATACCCCTAACGACAAA 1242
 Db |||||

RESULT 8
 ABA05546
 ID ABA05546 standard; cDNA; 1254 BP.
 AC
 AC ABA05546;
 XX
 XX
 DT 26-FEB-2002 (first entry)

Streptokinase cDNA.
 Streptokinase; cerebroprotective; cardiact; gene therapy; fusion;
 vasodilator; thrombolytic; angina; myocardial infarction; stroke;
 gene therapy; maxadilan; ss.
 Unidentified.
 WO200185100-A2.
 15-NOV-2001.
 10-MAY-2001; 2001WO-US015209.
 11-MAY-2000; 2000US-00569920.
 (GHO) GEN HOSPITAL CORP.
 Reddy VB, Lerner E;
 WPI; 2002-062184/08.
 New fusion protein or conjugate, useful for treating unstable angina,
 acute myocardial infarction or stroke, comprises a vasodilator
 polypeptide and a thrombolytic polypeptide, or active fragments of the
 polypeptides.
 Example 1; Fig 2; 37pp; English.
 The invention relates to a fusion protein or a conjugate comprising a
 vasodilator polypeptide, or its active fragment, and a thrombolytic
 polypeptide or its active fragment. The protein is useful for treating a
 subject suffering from a partially or totally occluded blood vessel,
 causing unstable angina, acute myocardial infarction or stroke. The
 polynucleotide encoding the polypeptide is useful in gene therapy. The
 vasodilatory action of the protein allows for the use of lower doses of a
 thrombolytic while maintaining the clot dissolving effectiveness of the
 thrombolytic, and the use of lower doses of the thrombolytic reduces
 associated side effects. The present sequence is the streptokinase cDNA
 used in the construction of a Maxadilan-Streptokinase fusion protein.
 Maxadilan is a vasodilator peptide produced by the salivary gland of the
 New World sand fly

Query Match 69.6%; Score 1240.4; DB 6; Length 1254;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 538 ATTGCTGACCTGAGTGCTGCTAGACCGTCACTGTCACAAACAGCCAAATGTTGTT 597
 Db |||||
 QY 7 ATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTCACAAACAGCCAAATGTTGTT 66
 Db |||||
 QY 598 AGCGTGTCTGCTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 657
 Db |||||
 QY 67 AGCGTGTCTGCTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 126
 Db |||||
 QY 658 GATCTAATCATCAGACCTGCTCATGAGGAGAAAGACAGACGAGGCTTAAGTCCAAATCA 717
 Db |||||
 QY 127 GATCTAATCATCAGACCTGCTCATGAGGAGAAAGACAGACGAGGCTTAAGTCCAAATCA 186
 Db |||||
 QY 718 AAACCAATTTGCTACTGATGAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTACTA 777
 Db |||||
 QY 187 AAACCAATTTGCTACTGATGAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTACTA 246
 Db |||||
 QY 778 AAGCTATTCAAGAACAAATTCATCGCTTAACGTCACAGTAAAGCAGTACTTTGAGTCTC 837
 Db |||||
 QY 247 AAGCTATTCAAGAACAAATTCATCGCTTAACGTCACAGTAAAGCAGTACTTTGAGTCTC 306
 Db |||||
 QY 838 ATTGATTTTTCGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCCTACTTTGCTGAC 897
 Db |||||
 QY 307 ATTGATTTTTCGAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGTCCTACTTTGCTGAC 366
 Db |||||

QY 898 AAGATGGTTCGGTAACCTTGCAGACCCAACTGCTCCAGAAATTTTGTCTAAGCGGACAT 957
 Db |||||
 QY 367 AAGATGGTTCGGTAACCTTGCAGACCCAACTGCTCCAGAAATTTTGTCTAAGCGGACAT 426
 Db |||||
 QY 958 GTGCGGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 1017
 Db |||||
 QY 427 GTGCGGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 486
 QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 1077
 Db |||||
 QY 487 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAAA 546
 QY 1078 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137
 Db |||||
 QY 547 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCAATCTCAAGAAATTACTA 606
 QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1197
 Db |||||
 QY 607 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 666
 QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
 Db |||||
 QY 667 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 726
 QY 1258 TTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Db |||||
 QY 727 TTTACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 786
 QY 1318 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
 Db |||||
 QY 787 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 846
 QY 1378 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAATCTGTTCACCATCAATACGTT 1437
 Db |||||
 QY 847 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAATCTGTTCACCATCAATACGTT 906
 QY 1438 GATGTCGATACCAACCAATTTGCTTAAAGTGACAGCTCTTAACAGCTAGGACGTAAC 1497
 Db |||||
 QY 907 GATGTCGATACCAACCAATTTGCTTAAAGTGACAGCTCTTAACAGCTAGGACGTAAC 966
 QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTTAACTACTCTCAACAAATCTC 1557
 Db |||||
 QY 967 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTTAACTACTCTCAACAAATCTC 1026
 QY 1558 GATGCTTTGGTATTATGACATATACCTTAATCGSAAAGTAGAGGATATACGATGAC 1617
 Db |||||
 QY 1027 GATGCTTTGGTATTATGACATATACCTTAATCGSAAAGTAGAGGATATACGATGAC 1086
 QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTAT 1677
 Db |||||
 QY 1087 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTAT 1146
 QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAGTTTACAGCTACTCG 1737
 Db |||||
 QY 1147 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAGTTTACAGCTACTCG 1206
 QY 1738 CATTATACAGGACACCTATACCTGATAACCCCTTAACGACAAA 1779
 Db |||||
 QY 1207 CGTTATACAGGACACCTATACCTGATAACCCCTTAACGACAAA 1248

RESULT 9
 ID ABA05547
 XX ABA05547 standard; DNA; 8893 BP.
 AC ABA05547;
 XX
 DT 26-FEB-2002 (first entry)
 DE Maxadilan-streptokinase fusion protein plasmid pRYB3maxetk.
 XX
 KW Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;

KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;
 KW gene therapy; maxadilan; sand fly; plasmid; ds.
 XX
 OS Lutzomyia longipalpis.
 OS Unidentified.
 OS Synthetic.
 OS Chimeric.
 PN WO200185100-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US015209.
 XX
 PR 11-MAY-2000; 2000US-00569920.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Reddy VB, Lerner B;
 XX
 XX WPI; 2002-062184/08.
 DR
 XX
 CC The invention relates to a fusion protein or a conjugate comprising a
 CC vasodilator polypeptide, or its active fragment, and a thrombolytic
 CC polypeptide or its active fragment. The protein is useful for treating a
 CC subject suffering from a partially or totally occluded blood vessel,
 CC causing unstable angina, acute myocardial infarction or stroke. The
 CC polynucleotide encoding the polypeptide is useful in gene therapy. The
 CC vasodilator action of the protein allows for the use of lower doses of a
 CC thrombolytic while maintaining the clot dissolving effectiveness of the
 CC thrombolytic, and the use of lower doses of the thrombolytic reduces
 CC associated side effects. The present sequence is the plasmid used for the
 CC preparation of Maxadilan-Streptokinase fusion protein. It contains
 CC maxadilan cDNA, which encodes a vasodilator peptide produced by the
 CC salivary gland of the New World sand fly, and streptokinase cDNA
 XX
 SQ Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match 69.6%; Score 1240.4; DB 6; Length 8893;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGCTAGACGCTCCATCTGTCACACAGCCAAATGGTTGTT 597
 Db |||||
 QY 5935 ATTGCTGGACCTGAGTGGCTGCTAGACGCTCCATCTGTCACACAGCCAAATGGTTGTT 5994
 QY 598 AGCGTTCGTCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
 Db |||||
 QY 5995 AGCGTTCGTCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 6054
 QY 658 GATCTAAACATCAGACCTGCTCATGAGGAGAAAGACAGACGAAGGCTTAAGTCCAAATCA 717
 Db |||||
 QY 6055 GATCTAAACATCAGACCTGCTCATGAGGAGAAAGACAGACGAAGGCTTAAGTCCAAATCA 6114
 QY 718 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTACTA 777
 Db |||||
 QY 6115 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAAATTGAGAAAGCTGACTACTA 6174
 QY 778 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTACAGCAGCTACTTTGAGGTC 837
 Db |||||
 QY 6175 AAGGCTATTCAAGAACCAATTGATCGCTAAAGTCCACAGTACAGCAGCTACTTTGAGGTC 6234
 QY 838 ATTGATTTTGAACGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 897
 Db |||||
 QY 6235 ATTGATTTTGAACGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 6294

Db 301 ATTGATTTTCAAGCGATGCAACCAATTAATGATCGAAACGCAAGGCTACTTTGCTGAC 360
 Qy 898 AAGATGTTTCCGTACCTTGCAGCCACCTGTCACGAATTTTGTAAAGCGGACAT 957
 Db 361 AAGATGTTTCCGTACCTTGCAGCCACCTGTCACGAATTTTGTAAAGCGGACAT 420
 Qy 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 1017
 Db 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAATCTGTTGATGTG 480
 Qy 1018 GAATATACGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 1077
 Db 481 GAATATACGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAAA 540
 Qy 1078 GATACTAAGCTATTGAACACACTAGCTATCGGTGACCAATCACTCAAGATTACTA 1137
 Db 541 GATACTAAGCTATTGAACACACTAGCTATCGGTGACCAATCACTCAAGATTACTA 600
 Qy 1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGAC 1197
 Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGAC 660
 Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
 Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
 Qy 1258 TTTACTACCGTTGTTAAATTCGGGAACAGCTTATAGATCAATATAAATAATCTGGTCTG 1317
 Db 721 TTTACTACCGTTGTTAAATTCGGGAACAGCTTATAGATCAATATAAATAATCTGGTCTG 780
 Qy 1318 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
 Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 Qy 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 1437
 Db 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCAACCATCAATACGTT 900
 Qy 1438 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
 Db 901 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
 Qy 1498 TTAGACTTCAGAGATTTTACGATCCTCGTGATAAGGCTTAACTACTCTTCAACAATCTC 1557
 Db 961 TTAGACTTCAGAGATTTTACGATCCTCGTGATAAGGCTTAACTACTCTTCAACAATCTC 1020
 Qy 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAAGATGAC 1617
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAAGATGAC 1080
 Qy 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGATGCTAGCTAT 1677
 Db 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGATGCTAGCTAT 1140
 Qy 1678 CATTTAGCCTATGATTAAGATCGTTTATACCGAAGAGAACGAGAGTTTACAGCTACCTG 1737
 Db 1141 CATTTAGCCTATGATTAAGATCGTTTATACCGAAGAGAACGAGAGTTTACAGCTACCTG 1200
 Qy 1738 CGTTATACAGGACACCTTATACCTGATTAACCCCTTAAACGACAAA 1779
 Db 1201 CGTTATACAGGACACCTTATACCTGATTAACCCCTTAAACGACAAA 1242

RESULT 11

AAX16633
 ID AAX16633 standard; DNA; 1242 BP.
 XX
 AC AAX16633;
 XX
 DX
 XX
 XX
 DE Streptococcus equisimilis mutant streptokinase K59E encoding DNA.

XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance; ds.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 CDS 1..1242
 FT /*tag= a
 FT /transl_except= (pos:40..42,aa:Asn)
 FT /note= "no stop codon given"
 FT
 XX US5876999-A.
 PN 02-MAR-1999.
 XX 06-DEC-1995; 95US-00568393.
 XX 06-DEC-1995; 95US-00568393.
 XX (NASC-) NAT SCI COUNCIL.
 XX Wu H;
 XX WPI; 1999-189643/16.
 DR P-PSDB; AAX94665.
 XX Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 FT
 XX Claim 4; Col 11-14; 17pp; English.
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence encodes mutant SK K59E. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPIg) to
 CC plasmin (HPI), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen
 XX
 SQ Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;
 Query Match 69.3%; Score 1235.6; DB 2; Length 1242;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1238; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 538 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTGGTTCTT 597
 Db 1 ATTGCTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTGGTTCTT 60
 Qy 598 AGCGTTGCTGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
 Db 61 AGCGTTGCTGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 120
 Qy 658 GATCTAACATCACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
 Db 121 GATCTAACATCACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCGGAATCA 180
 Qy 718 AAACCATTTGCTACTGATGTCGCGCATGTGCATATAACTTGAGAAAGCTGACTTACTA 777
 Db 181 AAACCATTTGCTACTGATGTCGCGCATGTGCATATAACTTGAGAAAGCTGACTTACTA 240
 Qy 778 AAGGCTATTCAAGAACCAATTGATCGCTAAACCTGACAGTAAACGACGACTACTTTGAGGTC 837
 Db 241 AAGGCTATTCAAGAACCAATTGATCGCTAAACCTGACAGTAAACGACGACTACTTTGAGGTC 300

```
QY 838 ATTGATTTTGCAGCGGATGCAACATTACTGATCGAAACCGCAAGGCTACTTTGCTGAC 897
DB |||||
QY 301 ATTGATTTTGCAGCGGATGCAACATTACTGATCGAAACCGCAAGGCTACTTTGCTGAC 360
DB |||||
QY 898 AAAGATGGTTCGGTACCTTGCAGCCACCTGTCGAAGATTTTCTAAGCGGACAT 957
DB |||||
QY 361 AAAGATGGTTCGGTACCTTGCAGCCACCTGTCGAAGATTTTCTAAGCGGACAT 420
DB |||||
QY 958 GTGCGGCTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
DB |||||
QY 421 GTGCGGCTTAGACCATATAAGAAACCAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
DB |||||
QY 1018 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTTACAGCAGGCTCTCAAA 1077
DB |||||
QY 481 GAATATACCTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTTACAGCAGGCTCTCAAA 540
DB |||||
QY 1078 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137
DB |||||
QY 541 GATACCTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600
DB |||||
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
DB |||||
QY 601 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
DB |||||
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
DB |||||
QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
DB |||||
QY 1258 TTTACTACCGTGTAAATAACCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
DB |||||
QY 721 TTTACTACCGTGTAAATAACCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
DB |||||
QY 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1377
DB |||||
QY 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 840
DB |||||
QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 1437
DB |||||
QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 900
DB |||||
QY 1438 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1497
DB |||||
QY 901 GATGTCGATACCAACGAATGCTTAAAGAGTGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 960
DB |||||
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAACTACTCTACAAATCTC 1557
DB |||||
QY 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAACTACTCTACAAATCTC 1020
DB |||||
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTACGAGGATATCAGATGAC 1617
DB |||||
QY 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTACGAGGATATCAGATGAC 1080
DB |||||
QY 1618 ACCAACCGTATCATACCGTTTATATGCGGCAAGCGACCGAAGAGAGAGATGCTAGCTAT 1677
DB |||||
QY 1081 ACCAACCGTATCATACCGTTTATATGCGGCAAGCGACCGAAGAGAGAGATGCTAGCTAT 1140
DB |||||
QY 1678 CATTTAGCCCTATGATAAGATCGTTATACCGAAGAGAGAGAGATTTACAGCTACCTG 1737
DB |||||
QY 1141 CATTTAGCCCTATGATAAGATCGTTATACCGAAGAGAGAGAGATTTACAGCTACCTG 1200
DB |||||
QY 1738 CGTTATACAGGACACCTATACCTGATACCCCTAACGAGAAA 1779
DB |||||
QY 1201 CGTTATACAGGACACCTATACCTGATACCCCTAACGAGAAA 1242
DB |||||
```

RESULT 12

AAA37628

ID AAA37628 standard; DNA; 1327 BP.

XX AC

XX AAA37628;

XX AC

DT 15-SEP-2003 (revised)

```
DT 13-OCT-2000 (first entry)
XX Streptokinase-NTR gene.
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
XX plasminogen; human; fibronectin; thrombolytic therapy;
XX cardiovascular disorder; ss.
XX Streptococcus dysgalactiae subsp. equisimilis.
XX EP1024192-A2.
XX 02-AUG-2000.
XX 23-DEC-1999; 99EP-00310541.
XX 24-DEC-1998; 98IN-DE003825.
XX (COUL ) CSIR COUNCIL SCI IND RES.
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
XX Yadav M;
XX WPI; 2000-516032/47.
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX thrombolytic therapy comprises a streptokinase fused with fibrin binding
XX domains of human fibronectin.
XX Example 1; Fig 14; 58pp; English.
XX This sequence represents a streptokinase-NTR (SK-NTR) gene (where NTR
XX stands for N-terminally repaired with native sequence). The invention
XX relates to a hybrid plasminogen activator (PA) comprises a polypeptide
XX fusion between streptokinase (SK), which are capable of plasminogen (PG)
XX activation, and fibrin binding regions of human fibronectin, which are
XX from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
XX possesses the ability to bind with fibrin independently and also
XX characteristically retains a PG activation ability which becomes evident
XX only after a pronounced duration, or lag, after exposure of the PA to a
XX suitable animal or human PG. The hybrid streptokinase-fibrin binding
XX domain polypeptides are useful in thrombolytic therapy for various kinds
XX of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
XX as well as kinetics of plasminogen activation that are distinct from that
XX of natural streptokinase in being characterised by a temporary delay, or
XX lag of several minutes in the natural rate of the catalytic conversion of
XX plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
XX can bind tightly with fibrin in blood clots soon after introduction into
XX the vascular system without significantly activating the circulating
XX blood plasminogen to plasmin, thus aiding in the localisation of the
XX plasminogen activation process to the site of pathological thrombus. This
XX overcomes systemic plasminogen activation encountered during clinical use
XX of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX Sequence 1327 BP; 456 A; 279 C; 250 G; 342 T; 0 U; 0 Other;
```

Query Match 69.1%; Score 1232.2; DB 3; Length 1327;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1237; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 538 ATTGCTGCAGCTGAGTGGCTGCTAGACGGTCCATCTGTCAACACAGCCAAATGGTTGTT 597
DB |||||
QY 83 ATAGCTGGTCCCTGAATGGCTACTAGATCGTCTTCTGTAATAACAGCCAAATGGTTGTT 142
DB |||||
QY 598 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 657
DB |||||
QY 143 AGCGTTCGTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 202
DB |||||
QY 658 GATCTAACATCACGACCTGCTCATGAGGAGAAAGACAGACGAGGCTTAAGTCCAAAATCA 717
DB |||||
QY 203 GATCTAACATCACGACCTGCTCATGAGGAGAAAGACAGACGAGGCTTAAGTCCAAAATCA 262
DB |||||
QY 718 AAACATTTGCTACTGTAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTACTA 777
DB |||||
```

Db 263 AARCCATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTACTA 322
 Qy 778 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 837
 Db 323 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGGTC 382
 Qy 838 ATTGATTTTGCAGCGATGCAACATTTACTGATCGAAACGCGCAAGGCTCTACTTTGCTGAC 897
 Db 383 ATTGATTTTGCAGCGATGCAACATTTACTGATCGAAACGCGCAAGGCTCTACTTTGCTGAC 442
 Qy 898 AAGATGTTTGGTAACTTTGCGGACCCAACTGTCGAAATTTTTCGTAAGCGGACAT 957
 Db 443 AAGATGTTTGGTAACTTTGCGGACCCAACTGTCGAAATTTTTCGTAAGCGGACAT 502
 Qy 958 GTGCGGTTAGACCATATAAAGAAACCAATATCAAAACCAAGCGAATCTGTTGATGTG 1017
 Db 503 GTGCGGTTAGACCATATAAAGAAACCAATATCAAAACCAAGCGAATCTGTTGATGTG 562
 Qy 1018 GAATATCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTTCAGACGAGTCTCAAA 1077
 Db 563 GAATATCTGTACAGTTTACTCCCTTTAAACCCCTGATGACGATTTTTCAGACGAGTCTCAAA 622
 Qy 1078 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 1137
 Db 623 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 682
 Qy 1138 GCTCAAGCAAAAGCAATTTAAACAAACCCAGGCTATAGGATTTATGAACGTGAC 1197
 Db 683 GCTCAAGCAAAAGCAATTTAAACAAACCCAGGCTATAGGATTTATGAACGTGAC 742
 Qy 1198 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAGTGGATCAAGAG 1257
 Db 743 TCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAGTGGATCAAGAG 802
 Qy 1258 TTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Db 803 TTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 862
 Qy 1318 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 1377
 Db 863 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 922
 Qy 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACGTTCACCATCAATAGCTT 1437
 Db 923 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACGTTCACCATCAATAGCTT 982
 Qy 1438 GATGTCGATACCAACGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
 Db 983 GATGTCGATACCAACGAAATGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1042
 Qy 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTTAACTCTTCAACAAATCTC 1557
 Db 1043 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTTAACTCTTCAACAAATCTC 1102
 Qy 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATATACGATGAC 1617
 Db 1103 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATATACGATGAC 1162
 Qy 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
 Db 1163 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1222
 Qy 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAGAGGATTTTACAGTACCTG 1737
 Db 1223 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAGAGGATTTTACAGTACCTG 1282
 Qy 1738 CGTTATACAGGACACCTATACCTGATTAACCCCTTAAACGAAATATA 1782
 Db 1283 CGTTATACAGGACACCTATACCTGATTAACCCCTTAAACGAAATATA 1327

AAT77778

ID AAT77778 standard; cDNA; 2566 BP.

XX AC AAT77778;

XX AC AAT77778;

DT 17-OCT-2003 (revised)

DT 01-OCT-1997 (first entry)

XX Coding sequence for plasminogen-binding fragment of Streptokinase.

XX Plasminogen-binding fragment; streptokinase; degradation; MBP;

XX thrombolytic agent; blood clot; bolus; maltose-binding protein; ds.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX W09641883-A1.

XX 27-DEC-1996.

XX 07-JUN-1996; 96WO-US009640.

XX 09-JUN-1995; 95US-00488940.

XX (HARD) HARVARD COLLEGE.

XX Reed GL;

XX WPI; 1997-065469/06.

XX Modified forms of streptokinase resistant to enzymatic cleavage - useful as thrombolytic agents in treating thrombosis and in medical equipment.

XX Example 1; Page 22-23; 65pp; English.

CC This sequence encodes the wild type plasminogen-binding fragment of streptokinase. The protein fragment encoded by this sequence was used in the design of a modified streptokinase has an in vitro degradation rate at least 2 times slower than that of native streptokinase. Compounds containing modified streptokinases are specifically used as thrombolytic agents for dissolving blood clots in vivo in a mammal, preferably at a dose of 20000 U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 2566 BP; 825 A; 545 C; 469 G; 727 T; 0 U; 0 Other;

Query Match 69.0%; Score 1230.4; DB 2; Length 2566;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1243; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCGATCTGTCAACACAGCCAAATTTGGTTGT 596

Db 896 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCGATCTGTCAACACAGCCAAATTTGGTTGT 955

Qy 597 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 656

Db 956 TAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 1015

Qy 657 CGATCTAACATCACGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTAAAGTCCAAATC 716

Db 1016 CGATCTAACATCACGACCTGCTCAT--AGGAAGACAGAGCAAGGCTTAAAGTCCAAATC 1073

Qy 717 AAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTACT 776

Db 1074 AAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTGACTTACT 1133

Qy 777 AAAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAACGACGACTTTGAGGT 836

Db 1134 AAAGGCTATTCAAGAACAAATTTGATCGCTTAACGTCACAGTAACGACGACTTTGAGGT 1193

Qy 837 CATTTGATTTTGCAGGAGATGCAACCAATTTACTGATCGAAACGCGCAAGGCTTACTTTGCTGA 896

Db 1194 CATTTGATTTTGCAGGAGATGCAACCAATTTACTGATCGAAACGCGCAAGGCTTACTTTGCTGA 1253

QY 897 CAAAGATGGTTGGTAACTTGGCGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACA 956
 DB 1254 CAAAGATGGTTGGTAACTTGGCGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACA 1313
 QY 957 TGTGGCGGTTAGACCATATAAGAAACCAATCAAAACCAAGCGGAATCTGTTGATGT 1016
 DB 1314 TGTGGCGGTTAGACCATATAAGAAACCAATCAAAACCAAGCGGAATCTGTTGATGT 1373
 QY 1017 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTAGACAGGCTCTCAA 1076
 DB 1374 GGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTAGACAGGCTCTCAA 1433
 QY 1077 AGATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1136
 DB 1434 AGATCTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1493
 QY 1137 AGCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGA 1196
 DB 1494 AGCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGA 1553
 QY 1197 CTCCTCAATCGTCACATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1256
 DB 1554 CTCCTCAATCGTCACATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGA 1613
 QY 1257 GTTTACTTACCGTGTAAATAATCGGAACCAAGCTTATAGGATCAATAAAAAATCTGGTCT 1316
 DB 1614 GTTTACTTACCGTGTAAATAATCGGAACCAAGCTTATAGGATCAATAAAAAATCTGGTCT 1673
 QY 1317 GAATGAAGAAATAAACACACCTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAAGG 1376
 DB 1674 GAATGAAGAAATAAACACACCTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAAGG 1733
 QY 1377 GGAAGAAGCGGTATGATCCCTTTCATCGCAGCTCACTTGAACCTGTTCAACCATCAATACGT 1436
 DB 1734 GGAAGAAGCGGTATGATCCCTTTCATCGCAGCTCACTTGAACCTGTTCAACCATCAATACGT 1793
 QY 1437 TGATGTTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAA 1496
 DB 1794 TGATGTTCGATACCAACGAATTTGCTAAAGAGTGAGCAGCTCTTAACAGCTAGCGAAGCTAA 1853
 QY 1497 CTTAGACTTCAGAGATTTATACGATCCTCTGTATAGGCTAAACTACTCTACACCAATCT 1556
 DB 1854 CTTAGACTTCAGAGATTTATACGATCCTCTGTATAGGCTAAACTACTCTACACCAATCT 1913
 QY 1557 CGATGCTTTTGGTATTATGACACTATACCTTAACCTGGAAGAGTACAGATTAATCAAGATGA 1616
 DB 1914 CGATGCTTTTGGTATTATGACACTATACCTTAACCTGGAAGAGTACAGATTAATCAAGATGA 1973
 QY 1617 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 1676
 DB 1974 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTA 2033
 QY 1677 TCATTTAGCTTATGATTAAGATCGTTATACCGAAGAGAAACGAGAACTTTACAGCTACCT 1736
 DB 2034 TCATTTAGCTTATGATTAAGATCGTTATACCGAAGAGAAACGAGAACTTTACAGCTACCT 2093
 QY 1737 GGGTTATACAGGACACCTTATACCTGATTAACCTTACAGCAATAA 1782
 DB 2094 GGGTTATACAGGACACCTTATACCTGATTAACCTTACAGCAATAA 2139

RESULT 14

ID AAQ12162 standard; DNA; 1458 BP.

XX AAQ12162;

XX 25-MAR-2003 (revised)

DT 17-SEP-1991 (first entry)

XX Factor Xa-cleavable hirudin-1EGR-streptokinase gene.

XX Fusion protein; blood clotting; coagulation; fibrinolysis;

KW antithrombotic; thrombolysis; ss.
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1452 /*tag= a
 FT mat_peptide 1..195 /*tag= b
 FT misc_RNA 196..207 /label= hirudin HV-1
 FT /*tag= c
 FT /label= linker
 FT /note= "encodes factor Xa cleavage site"
 FT mat_peptide 208..1449 /*tag= d
 FT /label= streptokinase
 PN WO9109125-A.
 XX 27-JUN-1991.
 PD 07-DEC-1989; 89GB-00027722.
 PF 07-DEC-1989; 89GB-00027722.
 PR 07-DEC-1990; 90WO-GB001911.
 XX (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
 PA Dawson KM, Hunter MG, Czaplowski LG;
 XX WPI: 1991-208151/28.
 PI P-PSDB; AAR12885.
 DR Fusion protein cleavage by blood clotting enzyme - for prodn. of
 DR fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 PT Disclosure; Page 96; 115pp; English.
 XX The sequence of the synthetic hirudin HV-1 genes was designed based on
 CC the published amino acid sequence (Dodd J., et al FEBS Letters 165 180
 CC (1984)). The sequence of streptokinase was obtd. from PCR amplified
 CC chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
 CC used for the PCR were based on the published DNA sequence of S.
 CC equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
 CC 357-362 [1985]). The two sequences were used to construct an expression
 CC vector in which the hirudin gene is linked to the streptokinase gene via
 CC a linking sequence encoding a cleavage site for factor Xa. The factor Xa
 CC is present at the site of the target thrombus so the active agents are
 CC released specifically at the place where clot formation is occurring. See
 CC also AAQ12153-Q12156, AAQ12158-Q12161 and AAQ12490. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX Sequence 1458 BP; 491 A; 316 C; 290 G; 361 T; 0 U; 0 Other;

Query Match 68.7%; Score 1223.4; DB 2; Length 1458;
 Best Local Similarity 98.3%; Pred. No. 0;
 Matches 1236; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 526 ACCGATGTTGTTAGCTTGTGGACCTGAGTGGCTGTAGACCTCCATCTGTCAACACAGC 585
 DB 196 ATCGAAGGTAGAATTTGCTGGACCTGAGTGGCTGTAGACCTGCTGTCAACACAGC 255
 QY 586 CAATTGCTTTAGCTTGTCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTT 645
 DB 256 CAATTGCTTTAGCTTGTCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTT 315
 QY 646 TTTTTCGAAATTCGATCTAAACATCAGACCTCTCTCATGAGGAAAGACAGACAGAGGCTTA 705
 DB 316 TTTTTCGAAATTCGATCTAAACATCAGACCTCTCTCATGAGGAAAGACAGACAGAGGCTTA 375
 QY 706 AGTCCAAATCAAAACCATTTGCTGATAGTGGCGGATGTGCATATAAACTTGAGAAA 765

Db 376 AGTCCAAAACCAAACTTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAATA 435
 QY 766 GCTGACTTACTAAAGCGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAAGCAGAC 825
 Db 436 GCTGACTTACTAAAGCGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGTAAGCAGAC 495
 QY 826 TACTTTTGAGGTCAATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGCAAGGTC 885
 Db 496 TACTTTTGAGGTCAATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGCAAGGTC 555
 QY 886 TACTTTTGCTGACAAAGAGATGGTTGCGTAACCTTTGCGGACCCCAACCTGTGCCAAGAAATTTTG 945
 Db 556 TACTTTTGCTGACAAAGAGATGGTTGCGTAACCTTTGCGGACCCCAACCTGTGCCAAGAAATTTTG 615
 QY 946 CTAACGCGACATGTCGGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAA 1005
 Db 616 CTAACGCGACATGTCGGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAA 675
 QY 1006 TCTGTTGATGTGGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACA 1065
 Db 676 TCTGTTGATGTGGAATATATCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACA 735
 QY 1066 CCAGGTCTCAAGATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCT 1125
 Db 736 CCAGGTCTCAAGATATAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCT 795
 QY 1126 CAAGAAATTAAGCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATAGATT 1185
 Db 796 CAAGAAATTAAGCTCAAGCAACAAGCAATTTTAAACAAAAACCAACCCAGGCTATAGATT 855
 QY 1186 TATGAACGCTGCTCTCAATCTGCTCACTGATGCAATGATGATTTTCCGTACGATTTTACCA 1245
 Db 856 TATGAACGCTGCTCTCAATCTGCTCACTGATGCAATGATGATTTTCCGTACGATTTTACCA 915
 QY 1246 ATGGATCAAGAGTTTACTTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAA 1305
 Db 916 ATGGATCAAGAGTTTACTTTACCATGTCAAAATTCGGGAACAAGCTTATAGATCAATAA 975
 QY 1306 AAATCTGCTGAATGAAGAAATAAACAACACTGATCTCTGAGAATAATTACGTC 1365
 Db 976 AAATCTGCTGAATGAAGAAATAAACAACACTGATCTCTGAGAATAATTACGTC 1035
 QY 1366 CTTAAAAAGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAAC 1425
 Db 1036 CTTAAAAAGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAAC 1095
 QY 1426 ATCAATACGTTGATGTGATACCAACGAATTTGCTAAAAAGTCAGCAGCTCTTAAACAGCT 1485
 Db 1096 ATCAATACGTTGATGTCAACCAACGAATTTGCTAAAAAGTCAGCAGCTCTTAAACAGCT 1155
 QY 1486 AGCGAACGTAATTAAGCTTCAAGAGATTTATAGATCTCTGATAGGCTAAACTACTC 1545
 Db 1156 AGCGAACGTAATTAAGCTTCAAGAGATTTATAGATCTCTGATAGGCTAAACTACTC 1215
 QY 1546 TACAACATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGAT 1605
 Db 1216 TACAACATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGAT 1275
 QY 1606 AATCAGATGACCAACCGTATACACCGTTTATATGGCAAGCGACCCGGAAGGAG 1665
 Db 1276 AATCAGATGACCAACCGTATACACCGTTTATATGGCAAGCGACCCGGAAGGAG 1335
 QY 1666 AATGCTAGCTATCAATTTAGCTTATGATAAGATTCGTTATACCGAAGAACGAGAAGTT 1725
 Db 1336 AATGCTAGCTATCAATTTAGCTTATGATAAGATTCGTTATACCGAAGAACGAGAAGTT 1395
 QY 1726 TACAGCTACCTCGGTTATACAGGACACCTTATACCTGATAACCTTAAACGCAATAA 1782
 Db 1396 TACAGCTACCTCGGTTATACAGGACACCTTATACCTGATAACCTTAAACGCAATAA 1452

RESULT 15

ADM01294
 ID ADM01294 standard; DNA; 7057 BP.
 XX
 AC ADM01294;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Plasmid pVTPA-Bstrep DNA sequence.
 XX
 KW vector; plasmid; artificial intergene region; plant;
 KW transplasmatic angiosperm; agronomic property; stress resistance;
 KW rbcL gene; ds.
 XX
 OS Unidentified.
 XX
 FN WO2004029256-A2.
 XX
 PD 08-APR-2004.
 XX
 PP 15-SEP-2003; 2003WO-CU000009.
 XX
 PP 27-SEP-2002; 2002CU-00000208.
 XX
 PR (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.
 PA Selman-Housein Sosa G, Aguiar Cabeza E, Gonzalez Quintero ADC;
 PI Ramos Gonzalez O;
 PI
 XX WPI; 2004-316131/29.
 XX
 XX DNA vector for transformation and expression in plastids, useful e.g. for
 PT producing pharmaceutical proteins or improving agronomic properties, has
 PT gene inserted in artificial intergene region.
 PT
 PS Example 10; SEQ ID NO 26; 74pp; Spanish.
 XX
 CC The invention relates to a DNA vector (A) for stable transformation and
 CC expression of genes (I) in plastids, where (I) is inserted in an
 CC artificial intergene region (AIR) formed by combining two 5'-untranslated
 CC regions (5'-UTRs) of genes that transcribe in different directions and
 CC are derived from plants of different divisions or classes. (A) are used
 CC to produce transplasmatic angiosperms that have improved agronomic
 CC properties (e.g. resistance to biotic or abiotic stress) or express a
 CC very wide range of agricultural, veterinary, pharmaceutical, nutritional
 CC or industrial products, e.g. enzymes, vaccinating antigens, cytokines or
 CC immunoglobulins. Use of (A) eliminates the need for a transposon for gene
 CC insertion; inserted genes do not require promoters and terminators; and
 CC the structure of flanking sequences in (A) ensures universal
 CC applicability. Also any selection marker in (A) can be eliminated by
 CC homologous recombination. (A) provides efficient and stable expression of
 CC genes without causing any functional alterations. This sequence
 CC represents the plasmid pVTPA-Bstrep, an example of the vector of the
 CC invention.
 XX
 SQ Sequence 7057 BP; 2029 A; 1543 C; 1608 G; 1877 T; 0 U; 0 Other;
 Query Match 68.6%; Score 1222.6; DB 12; Length 7057;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTGGTT 597
 Db 2023 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTGGTT 2082
 QY 598 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 657
 Db 2083 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 2142
 QY 658 GATCTTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
 Db 2143 GACCTTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 2202
 QY 718 AACCATTTCGTACTGATAGTGGCGCGATGTCAATAACTTGAGAAAGCTGACTTACTA 777

Job time : 1078.25 secs

```

|||||
Db 2203 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAANAAGCTGACTTACTA 2262
QY 778 AAGGCTATTCAAGAAACAATTGATCGCTAAAGGTCCACAGTAACGACGACTACTTTGAGGTC 837
Db 2263 AAGGCTATTCAAGAAACAATTGATCGCTAAAGGTCCACAGTAACGACGACTACTTTGAGGTC 2322
QY 838 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 897
Db 2323 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 2382
QY 898 AAGATGGTTTCGGTACCTTTGCCGACCCACCTGTCGAAGAAATTTTGGCTAAGCGGACAT 957
Db 2383 AAGATGGTTTCGGTACCTTTGCCGACCCACCTGTCGAAGAAATTTTGGCTAAGCGGACAT 2442
QY 958 GTGCGGTTTAGACCATATAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 2443 GTGCGGTTTAGACCATATAAGAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 2502
QY 1018 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCCTCAAA 1077
Db 2503 GAATATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCCTCAAA 2562
QY 1078 GATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 1137
Db 2563 GATACCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 2622
QY 1138 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGACGTGAC 1197
Db 2623 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGACGTGAC 2682
QY 1198 TCCTCAATCGTCACTCATGACCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db 2683 TCCTCAATCGTCACTCATGACCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 2742
QY 1258 TTTTACTTACCGTTTAAACCAAGGTAATAGGATCAATAAAAATCTGGTCTG 1317
Db 2743 TTTTACTTACCGTTTAAACCAAGGTAATAGGATCAATAAAAATCTGGTCTG 2802
QY 1318 AATGAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db 2803 AATGAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 2862
QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTTCAACATCAAAATACGTT 1437
Db 2863 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTTTCAACATCAAAATACGTT 2922
QY 1438 GATGTCGATACCAACGAATGCTTAAAGGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db 2923 GATGTCGATACCAACGAATGCTTAAAGGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 2982
QY 1498 TTAGACTTCAGAGATTTATAGATCTCTGATAGAGCTTAAAGCTTAACTCTTACAACAATCTC 1557
Db 2983 TTAGACTTCAGAGATTTATAGATCTCTGATAGAGCTTAAAGCTTAACTCTTACAACAATCTC 3042
QY 1558 GATGCTTTTGGTATTATGAGTACTTAACTGGAAGAGTACGAGTATACGATGAC 1617
Db 3043 GATGCTTTTGGTATTATGAGTACTTAACTGGAAGAGTACGAGTATACGATGAC 3102
QY 1618 ACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
Db 3103 ACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 3162
QY 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAACGAGAGTTTACAGCTACCTG 1737
Db 3163 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAACGAGAGTTTACAGCTACCTG 3222
QY 1738 CGTTATACAGGGACACTATACCTGATACCCCTAAACGACAAATAA 1782
Db 3223 CGTTATACAGGGACACTATACCTGATACCCCTAAACGACAAATAA 3267

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:24:19 ; Search time 314.619 Seconds
(without alignments)
10068.109 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tgcgttcacgttcgtcg.....ataacctaacgacaaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/PTUS COMB.seq.*
- 7: /cgn2_6/prodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/prodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1245.8	69.9	2385	3	US-09-211-542A-1
2	1240.4	69.6	1242	2	US-09-211-542A-5
3	1238.8	69.5	1242	2	US-08-568-393B-1
4	1235.6	69.3	1242	2	US-08-568-393B-2
5	1230.4	69.0	2566	2	US-08-488-940-19
6	1223.4	68.7	1458	2	US-07-854-596B-42
7	1222.6	68.6	1245	2	US-07-703-778D-1
8	1220.4	68.5	1335	2	US-07-854-596B-14
9	1220.4	68.5	1512	2	US-07-854-596B-27
10	1220.2	68.5	2589	2	US-07-854-596B-34
11	1219.4	68.4	1257	2	US-07-854-596B-25
12	1219.4	68.4	1317	2	US-07-854-596B-18
13	1216.4	68.3	1467	2	US-07-854-596B-46
14	1184	66.4	1209	3	US-09-374-038-11
15	1184	66.4	1209	3	US-09-658-179-11
16	1181.4	66.3	1245	3	US-09-374-038-12
17	1181.4	66.3	1245	3	US-09-658-179-12
18	1093.6	61.4	1122	3	US-09-374-038-13
19	1093.6	61.4	1122	3	US-09-658-179-13
20	1093.6	61.4	1158	3	US-09-374-038-14
21	1093.6	61.4	1158	3	US-09-658-179-14
22	1081	60.7	1119	2	US-07-854-596B-30
23	1080.8	60.7	2253	2	US-07-854-596B-39
24	1068	59.9	1068	3	US-09-211-542A-11

25	1065	59.8	2208	3	US-09-211-542A-3	Sequence 3, Appli
26	829	46.5	1262	9	5240845-3	Patent No. 5240845
27	826	46.4	1242	9	5240845-2	Patent No. 5240845
28	450	25.3	450	3	US-09-211-542A-13	Sequence 13, Appl
29	232.8	13.1	7679	3	US-09-220-132-38	Sequence 38, Appl
30	232.8	13.1	7680	3	US-09-023-655-1289	Sequence 1289, Ap
31	232.8	13.1	7680	6	PCT-US95-09819-6	Sequence 6, Appli
32	232.8	13.1	7803	2	US-08-551-356-1	Sequence 1, Appli
33	232.8	13.1	7803	6	PCT-US93-12687-1	Sequence 1, Appli
34	232.8	13.1	8044	3	US-09-566-921-135	Sequence 135, App
35	229.8	12.9	4811	3	US-09-813-718-13	Sequence 13, Appl
36	229.6	12.9	7705	2	US-08-259-885-16	Sequence 16, Appl
37	229.6	12.9	7705	2	US-08-826-885-16	Sequence 16, Appl
38	229.6	12.9	7705	9	5455158-2	Patent No. 5455158
39	228.8	12.8	4742	3	US-09-813-718-15	Sequence 15, Appl
40	228.4	12.8	4877	3	US-09-813-718-11	Sequence 11, Appl
41	228.4	12.8	5018	3	US-09-813-718-9	Sequence 9, Appli
42	228	12.8	4100	3	US-09-813-718-5	Sequence 5, Appli
43	228	12.8	4682	3	US-09-813-718-3	Sequence 3, Appli
44	228	12.8	4682	3	US-09-813-718-7	Sequence 7, Appli
45	228	12.8	5174	3	US-09-813-718-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-211-542A-1
; Sequence 1, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2385
US-09-211-542A-1

Query Match 69.9%; Score 1245.8; DB 3; Length 2385;

	Best Local Similarity	99.4%; Pred. No. 0;	Matches 1250; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy	523	TTCACCGATGTTTCGTATTGCTTGGAACCTCGAGTGGCTGTAGACCCTGCATCTGTCAACAAC	582
Dd	1129	TCCATCGAGGTAGATTGCTGGACCTGAGTGGCTGTAGACCCTCATCTGTCAACAAC	1188
Qy	583	AGCCAATTGTTGTTTAGGGTTCGTGTAATTTGAGGGGACGAATCAAGAATTAATGACTT	642
Dd	1189	AGCCAATTGTTGTTTAGGGTTCGTGTAATTTGAGGGGACGAATCAAGAATTAATGACTT	1248
Qy	643	AAATTTTTTGAAATTCGATCTAACAATCAGACCTGCTCATGGAGGAAGACAGACGAGGC	702
Dd	1249	AAATTTTTTGAAATTCGATCTAACAATCAGACCTGCTCATGGAGGAAGACAGACGAGGC	1308
Qy	703	TTAAGTCCAAAATCAAACCAATTTGCTACTGATAGTGGCGGATGTCAATATAAATTTGAG	762
Dd	1309	TTAAGTCCAAAATCAAACCAATTTGCTACTGATAGTGGCGGATGTCAATATAAATTTGAG	1368
Qy	763	AAAGCTGACTTACTAAAGGCTATTCAAAGAACAAATTGATCGCTAACGCCACAGTAACGAC	822
Dd	1369	AAAGCTGACTTACTAAAGGCTATTCAAAGAACAAATTGATCGCTAACGCCACAGTAACGAC	1428
Qy	823	GACTACTTTTGAGGTCATGATTTTGGCAAGCATCAACCAATTAATGATCGAAGCGCAAG	882
Dd	1429	GACTACTTTTGAGGTCATGATTTTGGCAAGCGATGCAACCAATTAATGATCGAAGCGCAAG	1488
Qy	883	GTCTACTTTGCTGCAAAAGATGGTTTCGGTAACCTTGCAGCCCCAACCTGTCCAAAGAAATTT	942
Dd	1489	GTCTACTTTGCTGCAAAAGATGGTTTCGGTAACCTTGCAGCCCCAACCTGTCCAAAGAAATTT	1548
Qy	943	TTGCTAAGCGGACATGTGCGGTTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCG	1002
Dd	1549	TTGCTAAGCGGACATGTGCGGTTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCG	1608
Qy	1003	AAATCTGTTGATGTGGAATATACGTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTC	1062
Dd	1609	AAATCTGTTGATGTGGAATATACGTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTC	1668
Qy	1063	AGACCAGTCTCAAAGATACTAAGCTATTGAAACACGTAGCTATCGGTGACACCATCACAC	1122
Dd	1669	AGACCAGTCTCAAAGATACTAAGCTATTGAAACACGTAGCTATCGGTGACACCATCACAC	1728
Qy	1123	TCTCAAGAAATTACTAGCTCAAGCAACAAGCATTTTAAACAAAACCAACCAGGCTATACG	1182
Dd	1729	TCTCAAGAAATTACTAGCTCAAGCAACAAGCATTTTAAACAAAACCAACCAGGCTATACG	1788
Qy	1183	ATTATGAAAGTGTACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTCGATTTTA	1242
Dd	1789	ATTATGAAAGTGTACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTCGATTTTA	1848
Qy	1243	CCAATGATCAAGAGTTTACTTACCGGTGTTAAABAATCGGGAACAAGCTTATAGATCAAT	1302
Dd	1849	CCAATGATCAAGAGTTTACTTACCGGTGTTAAABAATCGGGAACAAGCTTATAGATCAAT	1908
Qy	1303	AAAAAATCTGCTCTGAATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAATATTAC	1362
Dd	1909	AAAAAATCTGCTCTGAATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAATATTAC	1968
Qy	1363	GTCCCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTC	1422
Dd	1969	GTCCCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTC	2028
Qy	1423	ACCATCAAAATACGTTGATGTGATACCAACGAATTTGCTTAAAGTGAAGCAGCTCTTAAACA	1482
Dd	2029	ACCATCAAAATACGTTGATGTGATACCAACGAATTTGCTTAAAGTGAAGCAGCTCTTAAACA	2088
Qy	1483	GCTAGCGAAAGTAACTTACTTACTTACAGAGATTTTATACGATCCTCGTGATAAGGCTAAACTA	1542
Dd	2089	GCTAGCGAAAGTAACTTACTTACTTACAGAGATTTTATACGATCCTCGTGATAAGGCTAAACTA	2148
Qy	1543	CTCTPACAAATCTCGATGCTTTTGGTATTAATGACACTATACCTTTAACTGAAAGTAGAG	1602

```

Db      2149  CTCTACCAACAATCTCGATGCTTTTGGTGATATTATGGAAGTATACCTTTAATCTGGAAAAGTAGAG 2208
Qy      1603  GATAAATCAGCATGACACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGGAAGGA 1662
Db      2209  GATAAATCAGCATGACACCAACCGGTATCATAAACCGTTTATATGGCAAGCGACCCGGAAGGA 2268
Qy      1663  GAGAATCTAGCTATCATTTTAGCCTATGATAAAGATCGTTTATACCGNAGAAGAACGAGAA 1722
Db      2269  GAGAATGCTAGCTATCATTTTAGCCTATGATAAAGATCGTTTATACCGAAGAAGAACGAGAA 2328
Qy      1723  GTTTACAGCTACCTGCGCTTATACAGGACACCTTATACCTGATACCCCTAACGACAAA 1779
Db      2329  GTTTACAGCTACCTGCGCTTATACAGGACACCTTATACCTGATACCCCTAACGACAAA 2385

RESULT 2
US-09-211-542A-5
; Sequence 5, Application US/092111542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1242
US-09-211-542A-5

```

	Query Match	Best Local Similarity	Score 1240.4;	DB 3;	Length 1242;
	Matches 1241;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	538	ATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCCCAATTGGTTGTT	597		
Db	1	ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACACGCCCAATTAGTTGTT	60		
Qy	598	AGCGTTCCTGGTACTGTTTGGGGGACGAATCAGACAAATTAGTCTTAAATTTTTTGGAAATC	657		
Db	61	AGCGTTCCTGGTACTGTTTGGGGGACGAATCAGACAAATTAGTCTTAAATTTTTTGGAAATC	120		

QY 658 GATCTAATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGCGTTAAAGTCCAAATCA 717
Db 121 GATCTAATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGCGTTAAAGTCCAAATCA 180
QY 718 AAACCAATTTGCTAGTAGTGGCGGATGTCTACATAAACTTCAGAAAGCTGACTTACTA 777
Db 181 AAACCAATTTGCTAGTAGTGGCGGATGTCTACATAAACTTCAGAAAGCTGACTTACTA 240
QY 778 AAGGCTATTCAAGAAACAAATTTGATCGCTTAAGCTCCACAGTAACGACGACTACTTTGAGGTC 837
Db 241 AAGGCTATTCAAGAAACAAATTTGATCGCTTAAGCTCCACAGTAACGACGACTACTTTGAGGTC 300
QY 838 ATTGATTTTGAAGGATCGAACCATTAATCTGATCGAAACGGCAAGTCTACTTTGCTGAC 897
Db 301 ATTGATTTTGAAGGATCGAACCATTAATCTGATCGAAACGGCAAGTCTACTTTGCTGAC 360
QY 898 AAAGATGTTGGTAACTTTGCGGACCCCACTGTCGAAGAATTTTGTCTAAAGCGCAT 957
Db 361 AAAGATGTTGGTAACTTTGCGGACCCCACTGTCGAAGAATTTTGTCTAAAGCGCAT 420
QY 958 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAA 1077
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCTCAA 540
QY 1078 GATATCTAGCTATTGAACACTAGCTATCGGTGACACATCACTCTCAAGAAATTA 1137
Db 541 GATATCTAGCTATTGAACACTAGCTATCGGTGACACATCACTCTCAAGAAATTA 600
QY 1138 GCTCAAGCACAAAGCAATTTAAACCAAAACCAAGCGGTATACGATTTTATGAACGTGAC 1197
Db 601 GCTCAAGCACAAAGCAATTTAAACCAAAACCAAGCGGTATACGATTTTATGAACGTGAC 660
QY 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCTGACGATTTTACCAATGGAATCAAGAG 1257
Db 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCTGACGATTTTACCAATGGAATCAAGAG 720
QY 1258 TTTACTTACCGTTTAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGCTG 1317
Db 721 TTTACTTACCGTTTAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGCTG 780
QY 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATACGCTCTTAAAAAGGG 1377
Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATACGCTCTTAAAAAGGG 840
QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACCATCAATACGTT 1437
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACCATCAATACGTT 900
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAAACAGTACGCAACGTAAC 1497
Db 901 GATGTCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAAACAGTACGCAACGTAAC 960
QY 1498 TTAGACTTCAGAGATTTATACATCTCTGATAGGCTAAACTACTCTCAACAATCTC 1557
Db 961 TTAGACTTCAGAGATTTATACATCTCTGATAGGCTAAACTACTCTCAACAATCTC 1020
QY 1558 GATGCTTTTGGTATTTATGACATATACCTTTAACTGGAAGAGTAGAGGATATACGATGAC 1617
Db 1021 GATGCTTTTGGTATTTATGACATATACCTTTAACTGGAAGAGTAGAGGATATACGATGAC 1080
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGATGCTAGCTAT 1677
Db 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGAGAGATGCTAGCTAT 1140
QY 1678 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAGAAACGAGAGTTTACAGCTACCTG 1737
Db 1141 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAGAAACGAGAGTTTACAGCTACCTG 1200
QY 1738 CGTTATACAGGACACCTATACCTGATTAACCCCTTAACGACAA 1779

Db 1201 CGTTATACAGGACACCTATACCTGATATAACCTTAACGACAA 1242
RESULT 3
US-08-568-393B-1
; Sequence 1, Application US/08568393B
; Patent No. 5876999
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis H46A
; INDIVIDUAL ISOLATE: Malke, H., Roe, B., and Ferretti, J. J.;
; INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A" from Gene 34:357-362 (1985).
; CELL TYPE: Streptococcus equisimilis H46A
US-08-568-393B-1
Query Match 69.5%; Score 1238.8; DB 2; Length 1242;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 538 ATTGCTGGACCTCAGTGGCTGCTAGACCGCTCATCTGCAACACACCAATTTGTTGTT 597
Db 1 ATTGCTGGACCTCAGTGGCTGCTAGACCGCTCATCTGCAACACCAATTTGTTGTT 60
QY 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 657
Db 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTGAATC 120
QY 658 GATCTAATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGCGTTAAAGTCCAAATCA 717
Db 121 GATCTAATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGCGTTAAAGTCCAAATCA 180
QY 718 AAACCAATTTGCTAGTAGTGGCGGATGTCTACATAAACTTCAGAAAGCTGACTTACTA 777

```

Db 181 AAACATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA 240
QY 778 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTTGGGTC 837
Db 241 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTC 300
QY 838 ATTGATTTTGAAGCGATCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 897
Db 301 ATTGATTTTGAAGCGATCAACCAATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
QY 898 AAGATGTTTGGTACCTTTGCGGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 957
Db 361 AAGATGTTTGGTACCTTTGCGGACCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 420
QY 958 GTGCGGTTTGTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 421 GTGCGGTTTGTAGACCATATAAAGAAACCAATCAAAACCAAGCGAAATCTGTTGATGTG 480
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAGACGAGTCTCAAA 1077
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGTGATGACGATTTTCAGACGAGTCTCAAA 540
QY 1078 GATATAGCTATTGAAACACATAGCTATCGGTGACCAATCAATCTCAAGAAATTAATA 1137
Db 541 GATATAGCTATTGAAACACATAGCTATCGGTGACCAATCAATCTCAAGAAATTAATA 600
QY 1138 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 1197
Db 601 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCAAGAG 1257
Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCAAGAG 720
QY 1258 TTTACTACCGTGTAAAAATCGGAAACAAAGTTTATAGATCAATAAAAATCTGCTCTG 1317
Db 721 TTTACTACCGTGTAAAAATCGGAAACAAAGTTTATAGATCAATAAAAATCTGCTCTG 780
QY 1318 AATGAAGAAATAAACAACACTGACCTGATCTCGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCGAGAAATATTACGTCCTTAAAAAGGG 840
QY 1378 GAAAGGCGGTATGATCCCTTTGATCGCAGTCACTTCAAACTGTTCCACATCAATACGTT 1437
Db 841 GAAAGGCGGTATGATCCCTTTGATCGCAGTCACTTCAAACTGTTCCACATCAATACGTT 900
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db 901 GATGTCGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
QY 1498 TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTAAACTACTCTACAAATCTTC 1557
Db 961 TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTAAACTACTCTACAAATCTTC 1020
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTACGAGTAATACGATGAC 1617
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTACGAGTAATACGATGAC 1080
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1677
Db 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT 1140
QY 1678 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAAAGACGAGAAATTTACGCTACCTTG 1737
Db 1141 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAAAGACGAGAAATTTACGCTACCTTG 1200
QY 1738 CGTTATACGGGACACCTATACCTGATTAACCCCTTAACGACAAA 1779
Db 1201 CGTTATACGGGACACCTATACCTGATTAACCCCTTAACGACAAA 1242

```

```

; Sequence 2, Application US/08568393B
; Patent No. 5876999
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: SK-K59E
; LOCATION: DNA sequence No. 5876999174 and 175 have been changed
; LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
; LOCATION: from lys to Glu.
; OTHER INFORMATION:
; US-08-568-393B-2

```

```

Query Match 69.3%; Score 1235.6; DB 2; Length 1242;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1238; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCTTGGTGT 597
Db 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCTTGGTGT 60
QY 598 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 657
Db 61 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 120
QY 658 GATCTAACATCAGACCTCTCATGGAGGAAAGACAGACGAGCTTAAGTCCAAATCA 717
Db 121 GATCTAACATCAGACCTCTCATGGAGGAAAGACAGACGAGCTTAAGTCCGAAATCA 180
QY 718 AAACCATTTCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 777
Db 181 AAACCATTTCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTTACTA 240
QY 778 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTC 837
Db 241 AAGGCTATTCAAGAAACAAATTTGATCGCTAAAGTCCACAGTAACGACGACTACTTTGAGTC 300

```

```

QY 838 ATTGATTTTCAGCGATGCAACCACTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 897
Db 301 ATTGATTTTCAGCGATGCAACCACTTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
QY 898 AAAGATGGTTTCGGTAACTTTGCCGACCAACCTCTCCAGGAATTTTGGTAAAGCGGACAT 957
Db 361 AAAGATGGTTTCGGTAACTTTGCCGACCAACCTCTCCAGGAATTTTGGTAAAGCGGACAT 420
QY 958 GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db 421 GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAAGGTCTCAA 1077
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAAGGTCTCAA 540
QY 1078 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGATTTACTA 1137
Db 541 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCTCAAGATTTACTA 600
QY 1138 GCTCAAGCACAAGCATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGAC 1197
Db 601 GCTCAAGCACAAGCATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGAC 660
QY 1198 TCCTCAATCGTCACTCANGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 1257
Db 661 TCCTCAATCGTCACTCANGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 720
QY 1258 TTTTACTACCGTGTAAATTCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 721 TTTTACTACCGTGTAAATTCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 1318 AATGAAGAAATAAACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 781 AATGAAGAAATAAACACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTCAACCAATAGCTT 1437
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTCAACCAATAGCTT 900
QY 1438 GATGTCGATACCAACGAATGCTTAAAGTGAGAGAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db 901 GATGTCGATACCAACGAATGCTTAAAGTGAGAGAGCTCTTAAACAGCTAGCGAACGTAAC 960
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATGAGCTAACTTCTACCAACATCTC 1557
Db 961 TTAGACTTCAGAGATTTATACGATCCTCGTGATGAGCTAACTTCTACCAACATCTC 1020
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGMAAGTAGAGGATAATCAGATGAC 1617
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGMAAGTAGAGGATAATCAGATGAC 1080
QY 1618 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAAGAGAGAATGCTAGCTAT 1677
Db 1081 ACCAACCGTATACATACCGTTTATATGGCAAGCGACCCGAAGAGAGAATGCTAGCTAT 1140
QY 1678 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAACGAGTTTACAGTACCTG 1737
Db 1141 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAACGAGTTTACAGTACCTG 1200
QY 1738 CGTTATACAGGGACCTATACCTGATTAACCCCTAAACGACAAA 1779
Db 1201 CGTTATACAGGGACCTATACCTGATTAACCCCTAAACGACAAA 1242

```

RESULT 5

```

US-08-488-940-19
; Sequence 19, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Frazer, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-488-940-19

```

```

Query Match      69.0%; Score 1230.4; DB 2; Length 2566;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 537 TATTGCTGGAGCTGAGTGGCTGTAGACGGTCCATCTCTCAACACACGCAATTTGGTGT 596
Db 896 TATTGCTGGAGCTGAGTGGCTGTAGACGGTCCATCTCTCAACACACGCAATTTGGTGT 955
QY 597 TAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTTAAATTTTGAAT 656
Db 956 TAGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTTAAATTTTGAAT 1015
QY 657 CGATCTAACATCACGACCTGCTCATGGAGGAAAGACAGACAGCTTTAAGTCCAAAATC 716
Db 1016 CGATCTAACATCACGACCTGCTCAT--AGGAAAGACAGACAGCTTTAAGTCCAAAATC 1073
QY 717 AAAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACT 776
Db 1074 AAAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTTACT 1133
QY 777 AAAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGGT 836
Db 1134 AAAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGGT 1193
QY 837 CATTTGATTTTGAACGAGTACCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGA 896
Db 1194 CATTTGATTTTGAACGAGTACCAACCATTTACTGATCGAAACGGCAAGGTCTACTTTGCTGA 1253
QY 897 CAAAGATGGTTTCGGTAACTTTGCGGACCCAAACCTGTCCAAAGATTTTGTGTAAGCGGACA 956
Db 1254 CAAAGATGGTTTCGGTAACTTTGCGGACCCAAACCTGTCCAAAGATTTTGTGTAAGCGGACA 1313
QY 957 TGTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCAAGTCTGTGATGT 1016
Db 1314 TGTGCGGTTTAGACCATATAAAGAAACCAATACAAACCAAGCAAGTCTGTGATGT 1373
QY 1017 GGAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACCGAGTCTCAA 1076
Db 1374 GGAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACCGAGTCTCAA 1433

```


QY 1126 CAAGAATTACTAGTCAAGCACAAGAGCAATTTTAAACAAAACCCAGGCTATACGATT 1185
| | | | |
Db 796 CAAGAATTACTAGTCAAGCACAAGAGCAATTTTAAACAAAACCCATCCAGGCTATACGATT 855
| | | | |
QY 1186 TATGACGAGTACCTCCTCAATCGTCACTCATGACATGACATGATTTCCGTACGATTTTACCA 1245
| | | | |
Db 856 TATGACGAGTACCTCCTCAATCGTCACTCATGACATGACATGATTTCCGTACGATTTTACCA 915
| | | | |
QY 1246 ATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAA 1305
| | | | |
Db 916 ATGGATCAAGAGTTTACTTACCATGTCAAAATCGGAAACAGCTTATAGGATCAATAAA 975
| | | | |
QY 1306 AAATCTGGTCTGAATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTC 1365
| | | | |
Db 976 AAATCTGGTCTGAATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTC 1035
| | | | |
QY 1366 CTTAAATAAGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTCACC 1425
| | | | |
Db 1036 CTTAAATAAGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTCACC 1095
| | | | |
QY 1426 ATCAATACGTTGATGTCATACCAACGAATTTGCTAAATAAGTGACGAGCTCTTAACAGCT 1485
| | | | |
Db 1096 ATCAATACGTTGATGTCATACCAACGAATTTGCTAAATAAGTGACGAGCTCTTAACAGCT 1155
| | | | |
QY 1486 AGCGAAGTAACTTAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGGCTTAAACTACTC 1545
| | | | |
Db 1156 AGCGAAGTAACTTAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGGCTTAAACTACTC 1215
| | | | |
QY 1546 TACAACATCTCGATGCTTTTGGTATATAGGACTATACCTTAACTGAAAGTAGAGAT 1605
| | | | |
Db 1216 TACAACATCTCGATGCTTTTGGTATATAGGACTATACCTTAACTGAAAGTAGAGAT 1275
| | | | |
QY 1606 AATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCCGCGAGGAGAG 1665
| | | | |
Db 1276 AATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCCGCGAGGAGAG 1335
| | | | |
QY 1666 AATGCTAGTATCATTTAGCTATGATTAAGATGCTTTATACCGAAGAAACGAGAGATT 1725
| | | | |
Db 1336 AATGCTAGTATCATTTAGCTATGATTAAGATGCTTTATACCGAAGAAACGAGAGATT 1395
| | | | |
QY 1726 TACAGTACCTCGTTATACAGGACACCTATACCTGATACCTTAAACCTTAAACGACAAATAA 1782
| | | | |
Db 1396 TACAGTACCTCGTTATACAGGACACCTATACCTGATACCTTAAACCTTAAACGACAAATAA 1452
| | | | |

RESULT 7

US-07-703-778D-1.
; Sequence 1, Application US/07703778D
; Patent No. 5296366
; GENERAL INFORMATION:
; APPLICANT: Garcia, M. P. E. et al
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
; TITLE OF INVENTION: OF A GENE WHICH CODES FOR STREPTOKINASE, NUCLEOTIDE SEQUENCE
; TITLE OF INVENTION: OBTAINED, RECOMBINANT DNA AND TRANSFORMED MICROORGANISMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stanger, Michaelson, Spivak and Wallace, Esq.
; STREET: Parkway 109 Office Center, 328 Newman Springs Road,
; STREET: P. O. Box 8489
; CITY: Red Bank
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44Mb IBM compatible diskette
; COMPUTER: IBM PS/2 Model 80
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: Microsoft Word for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07703,778D
; FILING DATE: 19910522
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.
; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: Centro-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)530-6671
; TELEFAX: (908)530-6584
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1245 base pairs
; TYPE: NUCLEOTIDE WITH CORRESPONDING PROTEIN
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis from group C of Lanfield
; ORGANISM: definition
; IMMEDIATE SOURCE: ATCC-9542 strain
; FEATURE: from 1 to 1245 bp mature peptide
; OTHER INFORMATION:
; OTHER INFORMATION: Properties: Streptokinase gene
; OTHER INFORMATION: The gene product binds to human plasminogen
; OTHER INFORMATION: The gene product is an activator of human plasminogen
; US-07-703-778D-1

Query Match 68.6%; Score 1222.6; DB 2; Length 1245;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTTGTTGTT 597
| | | | |
Db 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTTGTTGTT 60
| | | | |
QY 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAATC 657
| | | | |
Db 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAATC 120
| | | | |
QY 658 GATCTAATCATCACCACCTGCTCATGGAGGAAGACAGCAAGCTTAAAGTCCAAATCA 717
| | | | |
Db 121 GACCTAATCATCACCACCTGCTCATGGAGGAAGACAGCAAGCTTAAAGTCCAAATCA 180
| | | | |
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTACTA 777
| | | | |
Db 181 AAACCATTTGCTACTGATAGTGGCGGATGTACATAAACTTGAGAAAGCTGACTACTA 240
| | | | |
QY 778 AAGCTATTCAGAAACAATTTGATCGCTAAACGTCACAGTACAGCACTACTTTGAGTC 837
| | | | |
Db 241 AAGGCTATTCAGAAACAATTTGATCGCTAAACGTCACAGTACAGCACTACTTTGAGTC 300
| | | | |
QY 838 ATTGATTTGCAAGCGATGCAACCATTAAGTATGATCGAAACGCAAGCTTCTTTGCTGAC 897
| | | | |
Db 301 ATTGATTTGCAAGCGATGCAACCATTAAGTATGATCGAAACGCAAGCTTCTTTGCTGAC 360
| | | | |
QY 898 AAGATGTTGCGTAACTCTTCCGACCAACCTGTCAGAGAAATTTTGTAAAGCGGACAT 957
| | | | |
Db 361 AAGATGTTGCGTAACTCTTCCGACCAACCTGTCAGAGAAATTTTGTAAAGCGGACAT 420
| | | | |
QY 958 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
| | | | |
Db 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
| | | | |
QY 1018 GAATATATCTGATGATTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCGCTCAAA 1077
| | | | |
Db 481 GAATATATCTGATGATTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCGCTCAAA 540
| | | | |
QY 1078 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
| | | | |
Db 541 GATACTAAGCTATTGAAACACATAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
| | | | |
QY 1138 GCTCAAGCACAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTCGAC 1197
| | | | |
Db 601 GCTCAAGCACAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTCGAC 660
| | | | |


```

QY 1198 TCCTCAATCGTCACTCATGACAAATGATGATTTTCCGTCAGATTTTACCAATGGATCAAGAG 1257
Db 661 TCCTCAATCGTCACTCATGACAAATGATGATTTTCCGTCAGATTTTACCAATGGATCAAGAG 720
QY 1258 TTTACTTACCGTGTGTTAAATAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 721 TTTACTTACCATGTCAAAAATCGGGAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
QY 1378 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAATCTGTTCAACATCAATACGTT 1437
Db 841 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAATCTGTTCAACATCAATACGTT 900
QY 1438 GATGTCGATACCAAGAAATGCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 1497
Db 901 GATGTCGATACCAAGAAATGCTTAAAGATGAGCAGCTCTTAAACAGCTAGCGAAGCTAAC 960
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATAGCTGATAGCTGATAGCTGATAGCTGAT 1557
Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATAGCTGATAGCTGATAGCTGATAGCTGAT 1020
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATGAGGATATACGATGAC 1617
Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATGAGGATATACGATGAC 1080
QY 1618 ACCAAGCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGAGAGAAATGCTAGCTAT 1677
Db 1081 ACCAAGCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGAGAGAAATGCTAGCTAT 1140
QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAGAGAGAAATGCTAGCTAT 1737
Db 1141 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAGAGAGAAATGCTAGCTAT 1200
QY 1738 CGTTATACAGGACACTATACCTGATAACCTTAAACGACAAATAA 1782
Db 1201 CGTTATACAGGACACTATACCTGATAACCTTAAACGACAAATAA 1245

```

RESULT 8

US-07-854-596B-14
; Sequence 14, Application US/07854596B
; Patent No. 5434073

GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000

```

; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1335 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1335
; OTHER INFORMATION: /note= "Streptokinase gene from S.
; OTHER INFORMATION: equisimilis"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1326
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 7..1326
; US-07-854-596B-14

Query Match 68.5%; Score 1220.4; DB 2; Length 1335;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1230; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 537 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTTGGTTGT 596
Db 84 TATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTTAGTTGT 143
QY 597 TAGCGTTGCTGGTACTCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTGAAT 656
Db 144 TAGCGTTGCTGGTACTCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTGAAT 203
QY 657 CGATCTAACATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATC 716
Db 204 TGACCTAACATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATC 263
QY 717 AAAACCAATTTGCTATGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACT 776
Db 264 AAAACCAATTTGCTATGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACT 323
QY 777 AAAGGCTATTCAAGAAACAATTTGATCGCTAACGCTCCACAGTAAACGACGACTACTTTGAGGT 836
Db 324 AAAGGCTATTCAAGAAACAATTTGATCGCTAACGCTCCACAGTAAACGACGACTACTTTGAGGT 383
QY 837 CATTTGATTTTCAAGCGATGCAACCAATTAATCTGATCGAAACGCGCAAGGCTTACTTTGCTGA 896
Db 384 CATTTGATTTTCAAGCGATGCAACCAATTAATCTGATCGAAACGCGCAAGGCTTACTTTGCTGA 443
QY 897 CAAAGATGGTTCCGTTAAACCTTTGCGGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACA 956
Db 444 CAAAGATGGTTCCGTTAAACCTTTGCGGACCCAACTGTCTCAAGAAATTTTGTCTAAGCGGACA 503
QY 957 TGTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGGAAATCTGTTGATGT 1016
Db 504 TGTGCGGTTAGACCATATAAAGAAACCAATAACAAACCAAGGAAATCTGTTGATGT 563
QY 1017 GGAATATCTGTACAGTTTACTCTCCCTTAAACCCCTGATGACGATTTCAAGACCGAGCTCTCAA 1076
Db 564 GGAATATCTGTACAGTTTACTCTCCCTTAAACCCCTGATGACGATTTCAAGACCGAGCTCTCAA 623
QY 1077 AGATATCTAAGCTATTGAAAAACATCTAGCTATCGGTGACACCATCACTCTCAAGAATTACT 1136
Db 624 AGATATCTAAGCTATTGAAAAACATCTAGCTATCGGTGACACCATCACTCTCAAGAATTACT 683
QY 1137 AGCTCAAGCAACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGACGCTGA 1196
Db 684 AGCTCAAGCAACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGACGCTGA 743
QY 1197 CTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTCAACATGATCAAGA 1256
Db 744 CTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTCAACATGATCAAGA 803

```

```

QY 1257 GTTACTTACCGTGTAAATAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCT 1316
Db 804 GTTACTTACCATGTCAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCT 863
QY 1317 GAATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAARAGG 1376
Db 864 GAATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAARAGG 923
QY 1377 GGAAGACCGGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTCCACCATCAATACGT 1436
Db 924 GGAAGACCGGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTCCACCATCAATACGT 983
QY 1437 TGATGTCGATACCAACGAATGCTTAAAGAGTGAAGCTCTTAACAGCTAGCGAACGTAA 1496
Db 984 TGATGTCGAACACCAACGAATGCTTAAAGAGTGAAGCTCTTAACAGCTAGCGAACGTAA 1043
QY 1497 CTTAGACTTCAGAGATTTATACGATCTCTGATGAAGCTTAAACTACTCTACCAATCT 1556
Db 1044 CTTAGACTTCAGAGATTTATACGATCTCTGATGAAGCTTAAACTACTCTACCAATCT 1103
QY 1557 CGATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTACAGGATATACAGTGA 1616
Db 1104 CGATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTACAGGATATACAGTGA 1163
QY 1617 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGGAAGGAGAGATGCTAGCTA 1676
Db 1164 CACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGGAAGGAGAGATGCTAGCTA 1223
QY 1677 TCATTTAGCCTATGATAAGATCGTTATACGGAAGAGACGAGAAAGTTTACAGTACCT 1736
Db 1224 TCATTTAGCCTATGATAAGATCGTTATACGGAAGAGACGAGAAAGTTTACAGTACCT 1283
QY 1737 GCGTTATACAGGACACCTATACCTGATAACCTTAACGACCAATAA 1782
Db 1284 GCGTTATACAGGACACCTATACCTGATAACCTTAACGACCAATAA 1329

```

RESULT 9

```

US-07-854-596B-27
; Sequence 27,-Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1234
; TELEFAX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 27:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1512 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1512
; OTHER INFORMATION: /note= "Streptokinase
; OTHER INFORMATION: fused to a yeast alpha-factor"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1503
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 7..1503
; US-07-854-596B-27

```

```

Query Match 68.5%; Score 1220.4; DB 2; Length 1512;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1233; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 529 GATGTTGCTATGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAA 588
Db 253 GATAAAGAAATGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAA 312
QY 589 TTGGTTGTTAGCGTTGCTGTACTCTTTGAGGGGAGCAATCAAGACATTAGTCTTAATTT 648
Db 313 TTAGTTGTTAGCGTTGCTGTACTCTTTGAGGGGAGCAATCAAGACATTAGTCTTAATTT 372
QY 649 TTTGAAATCGATCTAAACATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGGCTTAAGT 708
Db 373 TTTGAAATGACCTAAACATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGGCTTAAGT 432
QY 709 CCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCATGTCTCATATAAATTGAGAAAGCT 768
Db 433 CCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCATGTCTCATATAAATTGAGAAAGCT 492
QY 769 GACTTACTAAAGGCTATTCAAGAACAAATGATCGCTAAAGTCCACAGTAAACGACGACTAC 828
Db 493 GACTTACTAAAGGCTATTCAAGAACAAATGATCGCTAAAGTCCACAGTAAACGACGACTAC 552
QY 829 TTTGAGGTCAATGATTTTGGCAAGCGATGCAACCACTACTGATCGAAACGGCAAGGCTTAC 888
Db 553 TTTGAGGTCAATGATTTTGGCAAGCGATGCAACCACTACTGATCGAAACGGCAAGGCTTAC 612
QY 889 TTTGCTGACAAAGATGTTTCGGTAACTTGGCGCACCTTAAACCTGTCCAAGAAATTTTGCTA 948
Db 613 TTTGCTGACAAAGATGTTTCGGTAACTTGGCGCACCTTAAACCTGTCCAAGAAATTTTGCTA 672
QY 949 AGCGGACATGTGCGGTTAGACCATATATAAGAAACCAATACAAAACCAAGCGAAATCT 1008
Db 673 AGCGGACATGTGCGGTTAGACCATATATAAGAAACCAATACAAAACCAAGCGAAATCT 732
QY 1009 GTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGTGACGATTTGAGACA 1068
Db 733 GTTGATGTGGAATATATCTGTACAGTTTACTCCCTTAAACCTGTGACGATTTGAGACA 792
QY 1069 GGTCTCAAAAGATATCTAAAGTATTGAAACACATAGCTATCGGTGACACCATCATCTCAA 1128
Db 793 GGTCTCAAAAGATATCTAAAGTATTGAAACACATAGCTATCGGTGACACCATCATCTCAA 852
QY 1129 GAATTACTAGCTCAAGCAACAAAGCATTTTAAACAAAAACCCACCGGCTATACGATTTAT 1188
Db 853 GAATTACTAGCTCAAGCAACAAAGCATTTTAAACAAAAACCCACCGGCTATACGATTTAT 912
QY 1189 GAACGTGATCTCCTCAATCGTCACTCATGAATGACATTTTCCGTAACGATTTTACCAATG 1248
Db 913 GAACGTGATCTCCTCAATCGTCACTCATGAATGACATTTTCCGTAACGATTTTACCAATG 972
QY 1249 GATCAAGATTTTACTTACCGTGTAAATAATCGGAGCAAGCTTATAGATCAATAAAAAA 1308

```

```

Db 973 GATCAAGAGTTTACTTACATGTCAAAATCGGAAACAGCTTATGAGATCAATAAAAA 1032
QY 1309 TCTGGTCTCAATGAAGAAATAAACCAACACTGACTGATCTCTGAGAAATATACGTCCTT 1368
Db 1033 TCTGGTCTCAATGAAGAAATAAACCAACACTGACTGATCTCTGAGAAATATACGTCCTT 1092
QY 1369 AAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACATC 1428
Db 1093 AAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCAACATC 1152
QY 1429 AAATAGCTTGATGTGATACCAACGAATTTGCTTAAAGTGAGAGCTCTTAAACAGCTAGC 1488
Db 1153 AAATAGCTTGATGTGATACCAACGAATTTGCTTAAAGTGAGAGCTCTTAAACAGCTAGC 1212
QY 1489 GAACGTAACTTAGACTTCAGAGATTTATAGATCTCTCGTATAGGCTTAACTACTCTAC 1548
Db 1213 GAACGTAACTTAGACTTCAGAGATTTATAGATCTCTCGTATAGGCTTAACTACTCTAC 1272
QY 1549 AACAACTCTGATGCTTTTGGTATTTATGACTATACCTTAACTGGAAGTAGAGGATAAT 1608
Db 1273 AACAACTCTGATGCTTTTGGTATTTATGACTATACCTTAACTGGAAGTAGAGGATAAT 1332
QY 1609 CACGATGACCAACACCGTATCATAACCGTTTATATGCGGCAAGCGACCCGAAAGAGAGAAT 1668
Db 1333 CACGATGACCAACACCGTATCATAACCGTTTATATGCGGCAAGCGACCCGAAAGAGAGAAT 1392
QY 1669 GCTAGCTATCATTTAGCCTTATGATAAGATCGTTATACGGAAGCAAGCAAGAGTTTAC 1728
Db 1393 GCTAGCTATCATTTAGCCTTATGATAAGATCGTTATACGGAAGCAAGCAAGAGTTTAC 1452
QY 1729 AGCTACCTCGTTTATACAGGGACACCTATACCTTGATAACCTTAAACCAACAAATAA 1782
Db 1453 AGCTACCTCGTTTATACAGGGACACCTATACCTTGATAACCTTAAACCAACAAATAA 1506

```

RESULT 10

```

US-07-854-596B-34
; Sequence 34, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2589 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2589 /note=
; OTHER INFORMATION: "OmpAL-Streptokinase-streptokinase fusion linked
; OTHER INFORMATION: by thrombin-cleavable VELQGVVPRG"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..2580
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 4..2580
; US-07-854-596B-34

Query Match 68.5%; Score 1220.2; DB 2; Length 2589;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 534 TCGTATTGCTGGACCTGAGTGGCTGTAGACCCGTCATCTGTCAACACAGCCCAATTGGT 593
Db 1335 TGGAAATTGCTGGACCTGAGTGGCTGTAGACCCGTCATCTGTCAACACAGCCCAATTAGT 1394
QY 594 TGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGA 653
Db 1395 TGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGA 1454
QY 654 AATCGATCTAACATCACGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTTAAGTCCAAA 713
Db 1455 AATTGACCTTAACATCACGACCTGCTCATGGAGGAAGACAGAGCAAGGCTTTAAGTCCAAA 1514
QY 714 ATCAAAACCAATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGCTGACTT 773
Db 1515 ATCAAAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAAAAGCTGACTT 1574
QY 774 ACTAAAGGCTATTCAAGAACAAATTTGATCGCTCAAGTCCACAGTAACGACGACTACTTTGA 833
Db 1575 ACTAAAGGCTATTCAAGAACAAATTTGATCGCTCAAGTCCACAGTAACGACGACTACTTTGA 1634
QY 834 GGTCAATTGATTTTGCAGACGATGCAACCAATTTACTGTATCGAAACGGCAAGGCTCTACTTTCG 893
Db 1635 GGTCAATTGATTTTGCAGACGATGCAACCAATTTACTGTATCGAAACGGCAAGGCTCTACTTTCG 1694
QY 894 TGACAAAGATGGTTTCGGTAACTTTCGCGACCCCAACCTGTCCAGAAATTTTTCGTAAGCGG 953
Db 1695 TGACAAAGATGGTTTCGGTAACTTTCGCGACCCCAACCTGTCCAGAAATTTTTCGTAAGCGG 1754
QY 954 ACATGTGCGGTTAGACCATATATAAGAAACCAATACAAACCAAGCGGAATCTGTTGA 1013
Db 1755 ACATGTGCGGTTAGACCATATATAAGAAACCAATACAAACCAAGCGGAATCTGTTGA 1814
QY 1014 TGTGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCT 1073
Db 1815 TGTGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGCTCT 1874
QY 1074 CAAAGATACCTAAGCTATTGAAAAACATCTAGCTATCGGTGACACCATCATCTCAAGAAT 1133
Db 1875 CAAAGATACCTAAGCTATTGAAAAACATCTAGCTATCGGTGACACCATCATCTCAAGAAT 1934
QY 1134 ACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGCCAGGCTATACGATTTATGACG 1193
Db 1935 ACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGCCAGGCTATACGATTTATGACG 1994
QY 1194 TGACTCTCTCAATCGTCACTCATGACAATGACATTTTCGTCACGATTTTACCAATGGATCA 1253
Db 1995 TGACTCTCTCAATCGTCACTCATGACAATGACATTTTCGTCACGATTTTACCAATGGATCA 2054
QY 1254 AGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGG 1313
Db 2055 AGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTTATAGGATCAATAAAAAATCTGG 2114

```

QY 1314 TCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 1373
 Db 2115 TCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAA 2174
 QY 1374 AGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTTCCACATCAATA 1433
 Db 2175 AGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTTCCACATCAATA 2234
 QY 1434 CGTTGATGTCGATACCAACGAATGTCTAAAGTGGAGCAGCTCTTAACAGCTAGCGAACG 1493
 Db 2235 CGTTGATGTCGATACCAACGAATGTCTAAAGTGGAGCAGCTCTTAACAGCTAGCGAACG 2294
 QY 1494 TAACTTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAAGCTTAACTCTCAACAA 1553
 Db 2295 TAACTTAGACTTCAGAGATTATACGATCCCTCGTGATAAGGCTAAAGCTTAACTCTCAACAA 2354
 QY 1554 TCTCGATGCTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGATATACGA 1613
 Db 2355 TCTCGATGCTTTGGTATTATGGACTATACCTTAACCTGGAAGCTAGAGATATACGA 2414
 QY 1614 TGACACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAG 1673
 Db 2415 TGACACCAACCGGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAG 2474
 QY 1674 CTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAAGCGAGAAAGTTTACAGCTA 1733
 Db 2475 CTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAAGCGAGAAAGTTTACAGCTA 2534
 QY 1734 CTTGCTGTTATACAGGAGACCTATACCTGATTAACCTTAACCGAAGAAAGCGAGAAATAA 1782
 Db 2535 CTTGCTGTTATACAGGAGACCTATACCTGATTAACCTTAACCGAAGAAAGCGAGAAATAA 2583

RESULT 11

US-07-854-596B-25
 ; Sequence 25, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplewski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/854,596B
 ; FILING DATE: 03-JUN-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26,949
 ; REFERENCE/DOCKET NUMBER: 92,337
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 910-221-5317
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1257 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

MOLECULE TYPE: cdna
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..1257
 OTHER INFORMATION: /note= "Methionyl-streptokinase
 fusion protein"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4..1248
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 4..1248
 US-07-854-596B-25

Query Match 68.4%; Score 1219.4; DB 2; Length 1257;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 1229; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCTCATGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTC 597
 Db 7 ATTGCTGACCTGAGTGGCTGCTAGACCGTCTCATGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTC 66
 QY 598 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTC 657
 Db 67 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTC 126
 QY 658 GATCTAATCAATCAAGACCTGCTCATGAGGAGGAGCAAGAGCAAGAGCTTAAAGTCCAAATCA 717
 Db 127 GACCTAAATCAAGACCTGCTCATGAGGAGGAGCAAGAGCAAGAGCTTAAAGTCCAAATCA 186
 QY 718 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATATAAATTTGAGAAAGCTGACTTACTA 777
 Db 187 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATATAAATTTGAGAAAGCTGACTTACTA 246
 QY 778 AAGGCTATTCAAGAACCAATTGATCGCTAAACGCTCAAGTCCACAGTAAACGAGCAGCTACTTTGAGGTC 837
 Db 247 AAGGCTATTCAAGAACCAATTGATCGCTAAACGCTCAAGTCCACAGTAAACGAGCAGCTACTTTGAGGTC 306
 QY 838 ATTGATTTTGAAGCGATGCAACCAATTACTGATGCGTAAACGCGCAAGGCTACTTTGCTGAC 897
 Db 307 ATTGATTTTGAAGCGATGCAACCAATTACTGATGCGTAAACGCGCAAGGCTACTTTGCTGAC 366
 QY 898 AAAGATGTTTCGGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTTCGAAGCGGACAT 957
 Db 367 AAAGATGTTTCGGTAAACCTTCCGACCCCAACCTGTCGAAGAAATTTTTCGAAGCGGACAT 426
 QY 958 GTGCGCGTTAGACCATATAAAGAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 427 GTGCGCGTTAGACCATATAAAGAAAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 486
 QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGACGAGTCTCAAA 1077
 Db 487 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAAGACGAGTCTCAAA 546
 QY 1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACCATCAATCTCAAGAAATTA 1137
 Db 547 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACCATCAATCTCAAGAAATTA 606
 QY 1138 GCTCAAGCACAAAGCATTTTAAACAAAAACCAAGCGGCTATACGATTTTATGAACGTCAC 1197
 Db 607 GCTCAAGCACAAAGCATTTTAAACAAAAACCAAGCGGCTATACGATTTTATGAACGTCAC 666
 QY 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 1257
 Db 667 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGCATTTTACCAATGGATCAAGAG 726
 QY 1258 TTTACTTACCGTGTAAAAATTCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Db 727 TTTACTTACCATGTCAAAAATTCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 786
 QY 1318 AATCAAGAAATAAACAACACTGATGCTCTGAGAAATATTAGTCTTTAAAAAGGG 1377
 Db 787 AATCAAGAAATAAACAACACTGATGCTCTGAGAAATATTAGTCTTTAAAAAGGG 846

```

QY 1378 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 1437
Db |||||
QY 847 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 906
Db |||||
QY 1438 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAACAGTACGGAACGTAAC 1497
Db |||||
QY 907 GATGTCGATACCAACGAATGCTTAAAGTGAGCAGCTCTTAAACAGTACGGAACGTAAC 966
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATGAAGCTAACTCTTAAACGAATCTC 1557
Db |||||
QY 967 TTAGACTTCAGAGATTTATACGATCCTCGTGATGAAGCTAACTCTTAAACGAATCTC 1026
QY 1558 GATGCTTTTGGTATTATGCACTATACCTTAACTGGAAGCTAGAGGATAATCAGATGAC 1617
Db |||||
QY 1027 GATGCTTTTGGTATTATGCACTATACCTTAACTGGAAGCTAGAGGATAATCAGATGAC 1086
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGATGCTAGCTAT 1677
Db |||||
QY 1087 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGATGCTAGCTAT 1146
QY 1678 CATTTAGCTATGATTAAGATCGTTATACCGAAGAGAACGAGAGATTTACAGCTACCTG 1737
Db |||||
QY 1147 CATTTAGCTATGATTAAGATCGTTATACCGAAGAGAACGAGAGATTTACAGCTACCTG 1206
QY 1738 CGTTATACAGGACACCTATACCTGATTAACCTTAAACGACAAATAA 1782
Db |||||
QY 1207 CGTTATACAGGACACCTATACCTGATTAACCTTAAACGACAAATAA 1251

```

RESULT 12

```

US-07-854-596B-18.
; Sequence 18, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: 1..1317
; OTHER INFORMATION: /note= "OmpA fused to mature
; streptokinase gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1308
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 4..1308
; US-07-854-596B-18

Query Match 68.4%; Score 1219.4; DB 2; Length 1317;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGGTTGTT 597
Db |||||
QY 598 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 657
Db |||||
QY 127 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 186
QY 658 GATCTAACATCAGCACCTGCTCATGAGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 717
Db |||||
QY 187 GACCTAACATCAGCACCTGCTCATGAGAGGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 246
QY 718 AAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAACTTTGAGAAAGCTGACTACTA 777
Db |||||
QY 247 AAACCAATTTGCTACTGATAGTGGCGCGATGTCAATAACTTTGAGAAAGCTGACTACTA 306
QY 778 AAGGCTATTCAAGAAACAATTAAGTCAAGTAAAGTCCACAGTACGAGACTACTTTGAGTGC 837
Db |||||
QY 307 AAGGCTATTCAAGAAACAATTAAGTCAAGTAAAGTCCACAGTACGAGACTACTTTGAGTGC 366
QY 838 ATTGATTTTCAAGCGATGCAACCAATTAAGTCAAGTAAAGTCCACAGGCTCTACTTTGCTGAC 897
Db |||||
QY 367 ATTGATTTTCAAGCGATGCAACCAATTAAGTCAAGTAAAGTCCACAGGCTCTACTTTGCTGAC 426
QY 898 AAAGATGGTTGCGTAACTTTGCGGACCCAACTGTCCAAAGAAATTTTGTAAAGCGGACAT 957
Db |||||
QY 427 AAAGATGGTTGCGTAACTTTGCGGACCCAACTGTCCAAAGAAATTTTGTAAAGCGGACAT 486
QY 958 GTGCGCGTTAGACCAATATAAGAAACCAATATAAGAAACCAATATAAGAAACCAATATAAG 1017
Db |||||
QY 487 GTGCGCGTTAGACCAATATAAGAAACCAATATAAGAAACCAATATAAGAAACCAATATAAG 546
QY 1018 GAATATATCTGTACAGTTTACTTCCCTTAAACCTCTGATGACGATTTTCAAGACAGGCTCTCAA 1077
Db |||||
QY 547 GAATATATCTGTACAGTTTACTTCCCTTAAACCTCTGATGACGATTTTCAAGACAGGCTCTCAA 606
QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTACATCTCAAGAAATTA 1137
Db |||||
QY 607 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTACATCTCAAGAAATTA 666
QY 1138 GCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGCTATACGATTTTATGAACGCTGAC 1197
Db |||||
QY 667 GCTCAAGCACAAAGCAATTTTAAACAAAAACCAAGCTATACGATTTTATGAACGCTGAC 726
QY 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGAGTCAAGAG 1257
Db |||||
QY 727 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGGATTTTACCAATGAGTCAAGAG 786
QY 1258 TTTACTTACCGTGTAAAAAATCGGGAACAAAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db |||||
QY 787 TTTACTTACCATGTCAAAAATCGGGAACAAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 846
QY 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db |||||
QY 847 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 906
QY 1378 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 1437

```

Db 907 GAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACCATCAATACGTT 966
Qy 1438 GATGTCGATACCAACGAATTTCTAAAGATGAGCAGCTCTTTAAGCTAGCGAAGCTAAC 1497
Db 967 GATGTCACCAACGAATTTGCTAAAGAGGAGCAGCTCTTTAAGCTAGCGAAGCTAAC 1026
Qy 1498 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAGAGGCTAAACTACTCTACAAATCTC 1557
Db 1027 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAGAGGCTAAACTACTCTACAAATCTC 1086
Qy 1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGATGAGAGGATAATCAGCATGAC 1617
Db 1087 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGATGAGAGGATAATCAGCATGAC 1146
Qy 1618 ACCAACCGTATACAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1677
Db 1147 ACCAACCGTATACAAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1206
Qy 1678 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAAACGAGAAATTTACAGCTACCTG 1737
Db 1207 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAAACGAGAAATTTACAGCTACCTG 1266
Qy 1738 CGTTATACGGGACACCTATACCTGATAACCCCTAACGACAAATAA 1782
Db 1267 CGTTATACGGGACACCTATACCTGATAACCCCTAACGACAAATAA 1311

RESULT 13

US-07-854-596B-46
; Sequence 46, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1467 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1467
; OTHER INFORMATION: /note= "Streptokinase-hirudin
; OTHER INFORMATION: fusion linked by Factor Xa-cleavable IEGR"

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1449
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..1449
US-07-854-596B-46
Query Match 68.3%; Score 1216.4; DB 2; Length 1467;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1226; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 538 ATTGCTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAAACAGCCAAATGGTTGTT 597
Db 1 ATTGCTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAAACAGCCAAATGGTTGTT 60
Qy 598 AGCGTTGCTGCTACTGTTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTGGAAATC 657
Db 61 AGCGTTGCTGCTACTGTTGAGGGGCGAATCAAGACATTAGTCTTAAATTTTGGAAATC 120
Qy 658 GATCTTAACATCAGCACCTGCTCATGCGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
Db 121 GACCTTAACATCAGCACCTGCTCATGCGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
Qy 718 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTTGAGAAAGCTGACTTACTA 777
Db 181 AAACCATTTGCTACTGATAGTGGCGGATGCCAATAAATTTGAAAGAGCTGACTTACTA 240
Qy 778 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTGC 837
Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTGC 300
Qy 838 ATTGATTTTGCAGCGATGCAACCAATTTACTGATGAAACGGCAGGCTTACTTTGCTGAC 897
Db 301 ATTGATTTTGCAGCGATGCAACCAATTTACTGATGAAACGGCAGGCTTACTTTGCTGAC 360
Qy 898 AAAGATGTTTCGGTAAACCTTCGCCGACCCCAACCTGTCCAAAGAAATTTTTCGTAAGCGEACAT 957
Db 361 AAAGATGTTTCGGTAAACCTTCGCCGACCCCAACCTGTCCAAAGAAATTTTTCGTAAGCGEACAT 420
Qy 958 GTGCGCGTTAGACCATATAAAGAAAAAACAATATACAAAACCAAGCGAAATCTGTTGANTG 1017
Db 421 GTGCGCGTTAGACCATATAAAGAAAAAACAATATACAAAACCAAGCGAAATCTGTTGANTG 480
Qy 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACGAGGCTCAAA 1077
Db 481 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACGAGGCTCAAA 540
Qy 1078 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 1137
Db 541 GATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 600
Qy 1138 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCCAGGCTATAGCATTTATGAACGTCAC 1197
Db 601 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCCATCCAGGCTATACGATTTATGAACGTCAC 660
Qy 1198 TCCTCAATCGTCACATGACAAATGACATTTTCCGTAGATTTTACCAATGATCAAGAG 1257
Db 661 TCCTCAATCGTCACATGACAAATGACATTTTCCGTAGATTTTACCAATGATCAAGAG 720
Qy 1258 TTTTACTACCGTGTAAATAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 721 TTTTACTACCGTGTAAATAATCGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 780
Qy 1318 AATGAAGAAATAAAACACACCTGACCTGATCTCTGAGAAATATTAGTCTTAAAAAAGGG 1377
Db 781 AATGAAGAAATAAAACACACCTGACCTGATCTCTGAGAAATATTAGTCTTAAAAAAGGG 840
Qy 1378 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTTCAACATCAATACGTT 1437
Db 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAACCTGTTTCAACATCAATACGTT 900
Qy 1438 GATGTCGATACCAAGAAATTTGCTAAAAAGTGAGGAGCTCTTTAACAGCTAGCGAAACGTAAC 1497

[illegible]

```

RESULT 14
US-09-374-038-11
; Sequence 11, Application US/09374038
; Patent NO. 6309873
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Eldor Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6309873
; CURRENT APPLICATION NUMBER: US/09/374,038
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-374-038-11

```

Query Match	66.4%	Score 1184	DB 3	Length 1209
Best Local Similarity	98.8%	Pred. No. 0		
Matches 1193	Conservative 0	Mismatches 15	Indels 0	Gaps 0
Qy	575	TCAACAACAGCCAATTGGTTGTTAGCGTGTGCTGGTACTGTTGAGGGGACGAATCAAGACA	634	
Db	2	TGAACAACAGCCAATTAGTTGTTAGCGTGTGCTGTTGTTGAGGGGACGAATCAAGACA	61	
Qy	635	TTAGTCTTAAATTTTTTGAATCCGATCTAACATCAGCACCTGCTCATGAGGAGGAAGACAG	694	
Db	62	TTAGTCTTAAATTTTTTGAATTGACCTTAACATCAGCACCTGCTCATGAGGAGGAAGACAG	121	
Qy	695	AGCAAGGCTTAAAGTCCAAAATCAAACCATTTGCTACTGATAGTGGCGCGCATGTCCACATA	754	
Db	122	AGCAAGGCTTAAAGTCCAAATCAAACCATTTGCTACTGATAGTGGCGCGCATGCCACATA	181	
Qy	755	AACTTGAGAAGCTGACTTACTAAAGGCTATTTCAAAGAACAAATTGATCGCTTAACGTCACCA	814	
Db	182	AACTTGAAAAGCTGACTTACTAAAGGCTATTTCAAAGAACAAATTGATCGCTTAACGTCACCA	241	
Qy	815	GTAACGACGACTACTTTGAGGTCAATGATTTTGCACGGATGCACCAATTTACTGATCGAA	874	
Db	242	GTAACGACGACTACTTTGAGGTCAATGATTTTGCACGGATGCACCAATTTACTGATCGAA	301	

QY	875	ACGCCAAGGTCCTACTTTGGCTGTGACAAAGATGGTTTCGGTAAACCTTGGCCGACCAACCTGTGTCC	934
DB	302	ACGCCAAGGTCCTACTTTGGCTGTGACAAAGATGGTTTCGGTAAACCTTGGCCGACCAACCTGTGTCC	361
QY	935	AAGAAATTTTTGCTTAAGCGGACATGTGGCGTGTAGACCATATAAGAAAAACCAATACAAA	994
DB	362	AAGAAATTTTTGCTTAAGCGGACATGTGGCGTGTAGACCATATAAGAAAAACCAATACAAA	421
QY	995	ACCAAGCGAAATCTGTGTGATGTGGAAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG	1054
DB	422	ATCAAGCGAAATCTGTGTGATGTGGAAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG	481
QY	1055	ACGATTTTCAGACACAGGTCTCAAGATATCTAAGCTATTTGAAAAACACTAGCTATCGGTGACA	1114
DB	482	ACGATTTTCAGACACAGGTCTCAAGATATCTAAGCTATTTGAAAAACACTAGCTATCGGTGACA	541
QY	1115	CCATCACATCTCAAGAAATTAAGCTCAAGCACAAGAGCATTTTAAACAAAAACCAACCCAG	1174
DB	542	CCATCACATCTCAAGAAATTAAGCTCAAGCACAAGAGCATTTTAAACAAAAACCAACCCAG	601
QY	1175	GCTATACGATTTATGAACGCTGACTCCCTCAATCGTCACTCATGACAATGACATTTTCCGTA	1234
DB	602	GCTATACGATTTATGAACGCTGACTCCCTCAATCGTCACTCATGACAATGACATTTTCCGTA	661
QY	1235	CGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTGTAAATAATTCGGAAACAAGCTTATG	1294
DB	662	CGATTTTACCAATGGATCAAGAGTTTACTTACCGTGTGTAAATAATTCGGAAACAAGCTTATG	721
QY	1295	GGATCAATAAATAATCTGGTCTGAACGAAGAAATAACACACGTGACCTGATCTCTGAGA	1354
DB	722	AGATCAATAAATAATCTGGTCTGAATGAGAAATAAACACACGTGACCTGATCTCTGAGA	781
QY	1355	AATATTACGTCCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGA	1414
DB	782	AATATTACGTCCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGA	841
QY	1415	AACGTGTTCAACATCAAAATACGTTGATGTGCATACCAACGAATTTGCTAAAAAGTGAGCAGC	1474
DB	842	AACGTGTTCAACATCAAAATACGTTGATGTGCATACCAACGAATTTGCTAAAAAGTGAGCAGC	901
QY	1475	TCTTAAACAGCTTAGCGAAGCTTAACTTAGAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGG	1534
DB	902	TCTTAAACAGCTTAGCGAAGCTTAACTTAGAGACTTTCAGAGATTTTATACGATCCTCGTGATAAGG	961
QY	1535	CTAACTACTCTACAAACAATCTCGATGCTTTTGATTTATGGACTATACCTTAACTGGAA	1594
DB	962	CTAACTACTCTACAAACAATCTCGATGCTTTTGATTTATGGACTATACCTTAACTGGAA	1021
QY	1595	AAGTAGAGGATTAATCAGATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCGAC	1654
DB	1022	AAGTAGAGGATTAATCAGATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCGAC	1081
QY	1655	CCGAAGGAGAGATGCTAGCTATCATTTAGCCCTATGATAAGATCGTTATACCGAAGAG	1714
DB	1082	CCGAAGGAGAGATGCTAGCTATCATTTAGCCCTATGATAAGATCGTTATACCGAAGAG	1141
QY	1715	AACGAGAAGTTTACAGCTACTCTCGTTTATACAGGACACCTTATACCTGTATAACCCCTAAACG	1774
DB	1142	AACGAGAAGTTTACAGCTACTCTCGTTTATACAGGACACCTTATACCTGTATAACCCCTAAACG	1201
QY	1775	ACAAATAA	1782
DB	1202	ACAAATAA	1209

RESULT 15
US-09-658-179-11
; Sequence 11, Application US/09658179
; Patent No. 6413759
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalivo, Arlene Garcia

; APPLICANT: Menendez, Alina Seralena
 ; APPLICANT: Escalona, Elder Pupo
 ; APPLICANT: Masso, Julio Raul Fernandez
 ; APPLICANT: Griego, Martha De Jesus Gonzalez
 ; TITLE OF INVENTION: STREPTOKINASE MUTANTS
 ; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
 ; Patent No. 6413759
 ; CURRENT APPLICATION NUMBER: US/09/658,179
 ; CURRENT FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-09-658-179-11

Query Match 66.4%; Score 1184; DB 3; Length 1209;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1193; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy	575	TCACAAACAGCCAAATGGTGTAGCGTCTGCTGCTACTGTTGAGGGGACGAAATCAAGACA	634
Db	2	TGACACAGCCAAATAGTGTAGCGTCTGCTGCTACTGTTGAGGGGACGAAATCAAGACA	61
Qy	635	TTAGTCTTAAATTTTTTGAATTCGATCTAACATCAGCACCTGCTCATGAGGAAAGACAG	694
Db	62	TTAGTCTTAAATTTTTTGAATTTGACCTAACATCAGCACCTGCTCATGAGGAAAGACAG	121
Qy	695	AGCAAGGCTTAAGTCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTACATA	754
Db	122	AGCAAGGCTTAAGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTACATA	181
Qy	755	AACTTGAAGAAAGCTGACTTAAAGCTATTCAAGAACCAATTCATCGCTAAAGTCCACA	814
Db	182	AACTTGAAGAAAGCTGACTTAAAGGCTATTCAAGAACCAATTCATCGCTAAAGTCCACA	241
Qy	815	GTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCAATTTACTGTATCGAA	874
Db	242	GTAACGACGACTACTTTGAGGTCAATTTGCAAGCGATGCAACCAATTTACTGTATCGAA	301
Qy	875	ACGCAAGGTCTACTTTGCTGCAAAAGATGGTTGGTAACTTTGCCGACCCCAACCTGTCC	934
Db	302	ACGCAAGGTCTACTTTGCTGCAAAAGATGGTTGGTAACTTTGCCGACCCCAACCTGTCC	361
Qy	935	AGAATTTTGTCTAAGCGACATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAA	994
Db	362	AGAATTTTGTCTAAGCGACATGTGCGGTTAGACCATATAAAGAAAAACCAATACAAA	421
Qy	995	ACCAAGCGAAATCTGTTGATGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG	1054
Db	422	ATCAAGCGAAATCTGTTGATGGAATATCTGTACAGTTTACTCCCTTAAACCCCTGATG	481
Qy	1055	ACGATTTTCAGACGAGTCTCAAGATACCTAAGCTATTGAAACCACTAGCTATCGGTGACA	1114
Db	482	ACGATTTTCAGACGAGTCTCAAGATACCTAAGCTATTGAAACCACTAGCTATCGGTGACA	541
Qy	1115	CCATCACATCTCAAGAAATTTACTAGCTCAAGACAAAGCATTTTAAACAAAAACCCACAG	1174
Db	542	CCATCACATCTCAAGAAATTTACTAGCTCAAGACAAAGCATTTTAAACAAAAACCCACAG	601
Qy	1175	GCTATACGATTTTATGACGTGACTCTCAATCGTCACTCATGCAATGACATTTTCGGTA	1234
Db	602	GCTATACGATTTTATGACGTGACTCTCAATCGTCACTCATGCAATGACATTTTCGGTA	661
Qy	1235	CGATTTTACCAATGGATCAAGAGTTTACTTTACCGTGTAAATAATCGGAAACAGCTTATA	1294
Db	662	CGATTTTACCAATGGATCAAGAGTTTACTTTACCGTGTAAATAATCGGAAACAGCTTATG	721
Qy	1295	GGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACCACTGACCTGTCTCTGAGA	1354
Db	722	AGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAAACCACTGACCTGTCTCTGAGA	781

Search completed: February 1, 2006, 12:43:43
 Job time : 315.619 secs

Qy	1355	AATATTACGTCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCATCTTGA	1414
Db	782	AATATTACGTCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCATCTTGA	841
Qy	1415	AACGTTCACCATCAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAAAAGTGAAGC	1474
Db	842	AACGTTCACCATCAAAATACGTTGATGTCGATACCAACGAATTTGCTTAAAAAGTGAAGC	901
Qy	1475	TCTTAAACAGTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTATAGG	1534
Db	902	TCTTAAACAGTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTATAGG	961
Qy	1535	CTAAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGAATATACCTTTAACTGGAA	1594
Db	962	CTAAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGAATATACCTTTAACTGGAA	1021
Qy	1595	AAGTAGAGGATTAATCAGCATGACACCAACCGTATCATTAACCGTTTATATGGCAAGCGAC	1654
Db	1022	AAGTAGAGGATTAATCAGCATGACACCAACCGTATCATTAACCGTTTATATGGCAAGCGAC	1081
Qy	1655	CCGAAGGAGAGAAATGCTAGCTATCATTTAGCTTATGATAAAGATCGTTATATCCGAAGAG	1714
Db	1082	CCGAAGGAGAGAAATGCTAGCTATCATTTAGCTTATGATAAAGATCGTTATATCCGAAGAG	1141
Qy	1715	AACGAGAGTTTACAGCTACCTCGGTTATACAGGACACCTATACCTGATACCCCTAAGG	1774
Db	1142	AACGAGAGTTTACAGCTACCTCGGTTATACAGGACACCTATACCTGATACCCCTAAGG	1201
Qy	1775	ACAAATAA	1782
Db	1202	ACAAATAA	1209

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 17:06:30 ; Search time 1472.42 Seconds
(without alignments)
10008.068 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tcgcttcacgttcgctcgcg.....ataacctaacgacaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1782	100.0	1782	3	US-09-940-235-11
2	1684	94.5	2096	3	US-09-940-235-12
3	1245	69.9	1377	3	US-09-940-235-5
4	1243.4	69.8	1245	3	US-09-940-235-1
5	1232.2	69.1	1327	3	US-09-940-235-6
6	1150.2	64.5	1661	3	US-09-940-235-10
7	1134.2	63.6	1541	3	US-09-940-235-9
8	1081.2	60.7	1323	8	US-10-474-792-657
9	232.8	13.1	2127	6	US-10-210-120-49
10	232.8	13.1	2127	9	US-10-956-157-4288
11	232.8	13.1	2127	9	US-10-909-035-49
12	232.8	13.1	2443	7	US-10-741-601-70
13	232.8	13.1	2443	8	US-10-741-600-238
14	232.8	13.1	2488	7	US-10-741-601-75
15	232.8	13.1	2488	8	US-10-741-600-244
16	232.8	13.1	4295	6	US-10-144-194A-51
17	232.8	13.1	4295	8	US-10-491-566-51
18	232.8	13.1	6510	7	US-10-741-601-72
19	232.8	13.1	6510	8	US-10-741-600-241
20	232.8	13.1	6988	7	US-10-236-392-1
21	232.8	13.1	7361	7	US-10-236-392-3
22	232.8	13.1	7677	9	US-10-956-157-4995
23	232.8	13.1	7679	9	US-10-831-704-38

24	232.8	13.1	7680	3	US-09-964-824A-574	Sequence 574, Appl
25	232.8	13.1	7680	5	US-10-171-311-63	Sequence 63, Appl
26	232.8	13.1	7680	6	US-10-236-031B-69	Sequence 69, Appl
27	232.8	13.1	7680	6	US-10-374-979-75	Sequence 75, Appl
28	232.8	13.1	7680	7	US-10-182-936A-75	Sequence 75, Appl
29	232.8	13.1	7680	7	US-10-641-643-1289	Sequence 1289, Ap
30	232.8	13.1	7680	7	US-10-717-597-222	Sequence 222, App
31	232.8	13.1	7680	8	US-10-788-792-79	Sequence 79, App
32	232.8	13.1	7680	8	US-10-477-238A-654	Sequence 654, App
33	232.8	13.1	7680	8	US-10-680-287A-654	Sequence 654, App
34	232.8	13.1	7680	8	US-10-278-698-88	Sequence 88, Appl
35	232.8	13.1	7680	8	US-10-278-698-603	Sequence 603, App
36	232.8	13.1	7680	9	US-10-843-641A-5877	Sequence 5877, Ap
37	232.8	13.1	7680	9	US-10-477-173-654	Sequence 654, App
38	232.8	13.1	7680	9	US-10-852-335A-52	Sequence 52, Appl
39	232.8	13.1	7795	5	US-10-084-817-2	Sequence 2, Appli
40	232.8	13.1	7823	7	US-10-741-601-77	Sequence 77, Appl
41	232.8	13.1	7823	8	US-10-741-600-245	Sequence 245, App
42	232.8	13.1	7848	7	US-10-741-601-78	Sequence 78, Appl
43	232.8	13.1	7848	8	US-10-741-600-246	Sequence 246, App
44	232.8	13.1	7867	5	US-10-098-841-6	Sequence 6, Appli
45	232.8	13.1	7935	7	US-10-741-601-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-09-940-235-11
; Sequence 11, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-11

Query Match 100.0%; Score 1782; DB 3; Length 1782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCGCTTCACGTTTCGCTCGGTATCGGTGATTCATTCTGCTAACAGTAAGGCAACCCCGC	60
Db	1	TCGCTTCACGTTTCGCTCGGTATCGGTGATTCATTCTGCTAACAGTAAGGCAACCCCGC	60
Qy	61	CAGCCTAGCCGGGTCTTCAACGACGAGGACGACGATCATGCGCACCCCGTGGCAGGACCCA	120
Db	61	CAGCCTAGCCGGGTCTTCAACGACGAGGACGACGATCATGCGCACCCCGTGGCAGGACCCA	120
Qy	121	ACGCTGCCGAGATCTCGATCCCGGAAATTAATACACTCATATAGGGAGACCAAC	180

Db 121 |||||ACGTCGCCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGCGAGACCAAC 180
 Qy 181 GGTTCCTCCCTAGABATAATTTGTTTAACTTTAAGAGAGATATACCATGTCGCAAGC 240
 Db 181 GGTTCCTCCCTAGABATAATTTGTTTAACTTTAAGAGAGAGATATACCATGTCGCAAGC 240
 Qy 241 ACAACAGATTGTACCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCCTATGT 300
 Db 241 ACAACAGATTGTACCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCCTATGT 300
 Qy 301 GTCGAGAGAAACGTGGGAGAGAGCGAGCGACCATCACTTGCATCTTCTAGAAATAGATGC 360
 Db 301 GTCGAGAGAAACGTGGGAGAGAGCGAGCGACCATCACTTGCATCTTCTAGAAATAGATGC 360
 Qy 361 AACGATCAGGACACAGGACATCCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATAAT 420
 Db 361 AACGATCAGGACACAGGACATCCTATAGAAATTTGGAGACACCTGGAGCAAGAGGATAAT 420
 Qy 421 CGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGCGAGAGAGTGGAAAGTGTGAG 480
 Db 421 CGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGCGAGAGAGTGGAAAGTGTGAG 480
 Qy 481 AGGCAACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTCTTCCCGATGTTGGTAT 540
 Db 481 AGGCAACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTCTTCCCGATGTTGGTAT 540
 Qy 541 GCTGACCTGAGTGGTCTAGACCGTCCATCTGTGCAAGGCAACCGCCGAGAGTGGAAAGTGTGAG 600
 Db 541 GCTGACCTGAGTGGTCTAGACCGTCCATCTGTGCAAGGCAACCGCCGAGAGTGGAAAGTGTGAG 600
 Qy 601 GTTGCTGTGCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATCGAT 660
 Db 601 GTTGCTGTGCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATCGAT 660
 Qy 661 CTAAATCAACGACCTGCTCATGAGGAGAAACAGACGCAAGGCTTAAAGTCCAAATCAAAA 720
 Db 661 CTAAATCAACGACCTGCTCATGAGGAGAAACAGACGCAAGGCTTAAAGTCCAAATCAAAA 720
 Qy 721 CAAATTTGCTATGATAGTGGCGGATGTGCATATAAATTTGAGAAAGTGAATCTACTAAAG 780
 Db 721 CAAATTTGCTATGATAGTGGCGGATGTGCATATAAATTTGAGAAAGTGAATCTACTAAAG 780
 Qy 781 GCTATTCAAGAACTTGTATGCTTAAGTCCAGTAAACGACGACTACTTTGAGGTCATT 840
 Db 781 GCTATTCAAGAACTTGTATGCTTAAGTCCAGTAAACGACGACTACTTTGAGGTCATT 840
 Qy 841 GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGACAAA 900
 Db 841 GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGACAAA 900
 Qy 901 GATGTTGCGTAACTTGTGCGGACCCCACTGTCCAAAGATTTTGTGTAAGCGGACATGTG 960
 Db 901 GATGTTGCGTAACTTGTGCGGACCCCACTGTCCAAAGATTTTGTGTAAGCGGACATGTG 960
 Qy 961 CCGGTTAGCCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGGAA 1020
 Db 961 CCGGTTAGCCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGGAA 1020
 Qy 1021 TATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCCAGGCTTCAAGAT 1080
 Db 1021 TATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGCCAGGCTTCAAGAT 1080
 Qy 1081 ACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTAGCT 1140
 Db 1081 ACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTAGCT 1140
 Qy 1141 GAAGCAAAAGCATTTTAAACAAAACCCACCGCTATACGATTTTATGAACGTGACTCC 1200
 Db 1141 GAAGCAAAAGCATTTTAAACAAAACCCACCGCTATACGATTTTATGAACGTGACTCC 1200
 Qy 1201 TCAATCGTCACTCATCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTT 1260

Db 1201 TCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTT 1260
 Qy 1261 ACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTGAAT 1320
 Db 1261 ACTTACCGTGTAAAAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTGAAT 1320
 Qy 1321 GAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTTAAAAAAGGGGAA 1380
 Db 1321 GAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTTACGTCCTTTAAAAAAGGGGAA 1380
 Qy 1381 AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAATCTGTTCAACCATCAAAATAGCTTGAAT 1440
 Db 1381 AAGCCGTATGATCCCTTTGATCGCAGTCACTTGAATCTGTTCAACCATCAAAATAGCTTGAAT 1440
 Qy 1441 GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAACCTTA 1500
 Db 1441 GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAACCTTA 1500
 Qy 1501 GACTTTCAGAGATTTATACGATCTCTGTAAGGCTAAACTTACTTACCAACAATCTCGAT 1560
 Db 1501 GACTTTCAGAGATTTATACGATCTCTGTAAGGCTAAACTTACTTACCAACAATCTCGAT 1560
 Qy 1561 GCTTTTGGTATTTATGGAATATACCTTAACTGGAAGAGTAGAGGATAATCAGATGACACC 1620
 Db 1561 GCTTTTGGTATTTATGGAATATACCTTAACTGGAAGAGTAGAGGATAATCAGATGACACC 1620
 Qy 1621 AACCGTATCATACCGTTTATATGGCAAGCGACCGAAGGAGAGATGCTAGCTATCAT 1680
 Db 1621 AACCGTATCATACCGTTTATATGGCAAGCGACCGAAGGAGAGATGCTAGCTATCAT 1680
 Qy 1681 TTAGCTTATGATTAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTCTGCT 1740
 Db 1681 TTAGCTTATGATTAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTCTGCT 1740
 Qy 1741 TATACAGGACACCTTATACCTGATTAACCTTAAACGACAAATAA 1782
 Db 1741 TATACAGGACACCTTATACCTGATTAACCTTAAACGACAAATAA 1782

RESULT 2

US-09-940-235-12
 ; Sequence 12, Application US/09940235
 ; Publication No. US2003005921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 2096
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hybrid cassette
 US-09-940-235-12

Query Match 94.5%; Score 1684; DB 3; Length 2096;

Best Local Similarity 99.7%; Pred. No. 0; Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
QY	1	TCGCTTCACGTTTCGCTCGCTATCGGTGATTCATTCCTGCTAACAGTAAGCAACCCCGC	60						
Db	51	TCGCTTCACGTTTCGCTCGCTATCGGTGATTCATTCCTGCTAACAGTAAGCAACCCCGC	110						
QY	61	CAGCTAGCGGGTCTCAACGACAGGACACGATCATGCGCACCCGTCGCGCAGACCCA	120						
Db	111	CAGCTAGCGGGTCTCAACGACAGGACACGATCATGCGCACCCGTCGCGCAGACCCA	170						
QY	121	ACGCTCCCGAGATCTCGATCCCGCGAATTAATACGACTCATATAGGAGACACAC	180						
Db	171	ACGCTCCCGAGATCTCGATCCCGCGAATTAATACGACTCATATAGGAGACACAC	230						
QY	181	GGTTTCCCTCTAGAAATAATTTTGTAACTTTAAGAAAGGAGATATACCATGTCGAAGC	240						
Db	231	GGTTTCCCTCTAGAAATAATTTTGTAACTTTAAGAAAGGAGATATACCATGTCGAAGC	290						
QY	241	ACAACAGATTGTACCATAGCTGAGAAAGTGTTTTGATCATGCTGCGGACTTCCTATGT	300						
Db	291	ACAACAGATTGTACCATAGCTGAGAAAGTGTTTTGATCATGCTGCGGACTTCCTATGT	350						
QY	301	GGTCGAGAAAACGTGGGAGAGCGGACGATCACTTGCATCTTGAATAAGATGTC	360						
Db	351	GGTCGAGAAAACGTGGGAGAGCGGACGATCACTTGCATCTTGAATAAGATGTC	410						
QY	361	AACGATCAGGACACAAAGGACATCTATAGAAATGGAGACACCTGGACCAAGAGATAAT	420						
Db	411	AACGATCAGGACACAAAGGACATCTATAGAAATGGAGACACCTGGACCAAGAGATAAT	470						
QY	421	CGAGGAAAACCTGCTCAGATGTCATCTGCACAGGCAACGGCCGAGGAGTGAAGTGTAG	480						
Db	471	CGAGGAAAACCTGCTCAGATGTCATCTGCACAGGCAACGGCCGAGGAGTGAAGTGTAG	530						
QY	481	AGGCACACTCTGTGTCAGACCAATCGAGCGGATCTGGCCCTTCACCGATGTCGTATT	540						
Db	531	AGGCACACTCTGTGTCAGACCAATCGAGCGGATCTGGCCCTTCACCGATGTCGTATT	590						
QY	541	GCTGGACCTGAGTGGTGTAGACCTGCATCTGTCAACAAACGCAATGTTGTTTAC	600						
Db	591	GCTGGACCTGAGTGGTGTAGACCTGCATCTGTCAACAAACGCAATGTTGTTTAC	650						
QY	601	GTTCGTGTACTGTTTCAGGGGAGCAATCAAGACATAGTCTTAAATTTTGAATCGAT	660						
Db	651	GTTCGTGTACTGTTTCAGGGGAGCAATCAAGACATAGTCTTAAATTTTGAATCGAT	710						
QY	661	CTAACATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAGTCCAAATCAAAA	720						
Db	711	CTAACATCAGACCTGCTCATGAGGAAAGACAGACGAGCTTAAGTCCAAATCAAAA	770						
QY	721	CAATTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTAAAG	780						
Db	771	CAATTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTAAAG	830						
QY	781	GCTATTCAAGACAAATTGATCGCTACGTCACAGTAAACGACCTACTTTGAGTCAAT	840						
Db	831	GCTATTCAAGACAAATTGATCGCTACGTCACAGTAAACGACCTACTTTGAGTCAAT	890						
QY	841	GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGTCACAAA	900						
Db	891	GATTTTGCAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGTCACAAA	950						
QY	901	GATGTTTCGGTAACTTTCGCGACCCCACTGTCACAGAAATTTTTCGTAAGCGGACATGTG	960						
Db	951	GATGTTTCGGTAACTTTCGCGACCCCACTGTCACAGAAATTTTTCGTAAGCGGACATGTG	1010						
QY	961	CGCGTTAGACCATATAAGAAACCAATACAAAACGACGAAATCTGTTGATGTGGA	1020						
Db	1011	CGCGTTAGACCATATAAGAAACCAATACAAAACGACGAAATCTGTTGATGTGGA	1070						
QY	1021	TATACGTACGTTTACTCCCTTAACCTTAAACCTGATGAGATTTTCAGCCAGCTCTCAAGAT	1080						

Db	1071	TATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAAGAT	1130
QY	1081	ACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTAGCT	1140
Db	1131	ACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTAGCT	1190
QY	1141	CAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGACTCC	1200
Db	1191	CAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGACTCC	1250
QY	1201	TCAATCGTCACTCATGACAATGACATTTTCGTGACATTTTACCAATGATCAAGATTT	1260
Db	1251	TCAATCGTCACTCATGACAATGACATTTTCGTGACATTTTACCAATGATCAAGATTT	1310
QY	1261	ACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTGAAT	1320
Db	1311	ACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTGAAT	1370
QY	1321	GAAGAAATAAACACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGGAA	1380
Db	1371	GAAGAAATAAACACACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAAGGGAA	1430
QY	1381	AAGCGGTATGATCCCTTTTGATCGGAGTCATCTTGAAACTGTTTCAACATCAATCTGAT	1440
Db	1431	AAGCGGTATGATCCCTTTTGATCGGAGTCATCTTGAAACTGTTTCAACATCAATCTGAT	1490
QY	1441	GTGATACCAAGCAATTCCTAAAGTGAGAGCTCTTAAACAGCTAGCGAGTAACTTA	1500
Db	1491	GTGATACCAAGCAATTCCTAAAGTGAGAGCTCTTAAACAGCTAGCGAGTAACTTA	1550
QY	1501	GACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAACTACTCTACAACTCTCGAT	1560
Db	1551	GACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAACTACTCTACAACTCTCGAT	1610
QY	1561	GCTTTTGTATTTAGGACTATACCTTAACTGAAAAGTAGAGGATAATCAGATGACACC	1620
Db	1611	GCTTTTGTATTTAGGACTATACCTTAACTGAAAAGTAGAGGATAATCAGATGACACC	1670
QY	1621	AACGTTATCAACCGTTTATATGGGCAAGGACCCGAGAGAGAGATGCTAGCTATCAT	1680
Db	1671	AACGTTATCAACCGTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTATCAT	1730
QY	1681	TTAGCTTATGAT 1692	
Db	1731	TTAGCTTGTGT 1742	

RESULT 3
US-09-940-235-5
; Sequence 5, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 05/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5

```

; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5

Query Match      69.9%; Score 1245; DB 3; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCATCTGTCAACACAGCCAAATGGTGTGTT 597
Db |||
Qy 133 ATTGCTGGACCTGAGTGGCTGTAGACCGTCATCTGTCAACACAGCCAAATGGTGTGTT 192
Db |||
Qy 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
Db |||
Qy 193 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 252
Db |||
Qy 658 GATCTAAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 717
Db |||
Qy 253 GATCTAAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 312
Db |||
Qy 718 AAACATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTACTA 777
Db |||
Qy 313 AAACATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTACTA 372
Db |||
Qy 778 AAGGCTATTCAAGAACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGTTC 837
Db |||
Qy 373 AAGGCTATTCAAGAACAAATTTGATCGCTTAAGTCCACAGTAACGACGACTACTTTGAGTTC 432
Db |||
Qy 838 ATTGATTTTTCAAGCGATGCAACATTTACTGATCGAAACCGGCAAGGCTCTACTTTGCTGAC 897
Db |||
Qy 433 ATTGATTTTTCAAGCGATGCAACATTTACTGATCGAAACCGGCAAGGCTCTACTTTGCTGAC 492
Db |||
Qy 898 AAGATGTTTGGTAACTTTGCGGACCAACCTGTCTCAAGAAATTTTTGCTTAAAGCGACAT 957
Db |||
Qy 493 AAGATGTTTGGTAACTTTGCGGACCAACCTGTCTCAAGAAATTTTTGCTTAAAGCGACAT 552
Db |||
Qy 958 GTGCGGTTAGACCATATAAAGAAACCAATATCAAAACCAAGGAAATCTGTTGATGTG 1017
Db |||
Qy 553 GTGCGGTTAGACCATATAAAGAAACCAATATCAAAACCAAGGAAATCTGTTGATGTG 612
Db |||
Qy 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA 1077
Db |||
Qy 613 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTACAGCAGGCTCTCAAA 672
Db |||
Qy 1078 GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 1137
Db |||
Qy 673 GATACTAAGCTATTGAAACACCTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 732
Db |||
Qy 1138 GCTCAAGCAAAAGCAATTTAAACAAACCCAGGCTATACGATTTATGAACGTGAC 1197
Db |||
Qy 733 GCTCAAGCAAAAGCAATTTAAACAAACCCAGGCTATACGATTTATGAACGTGAC 792
Db |||
Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db |||
Qy 793 TCCTCAATCGTCACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 852
Db |||
Qy 1258 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db |||
Qy 853 TTTACTTACCGTGTAAAAATCGGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 912
Db |||
Qy 1318 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db |||
Qy 913 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 972
Db |||
Qy 1378 GAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAACTGTTCCCAATCAATACGTT 1437
Db |||
Qy 973 GAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAACTGTTCCCAATCAATACGTT 1032
Db |||
Qy 1438 GATGTCGATACCAACGAATTTGCTTAAAAAGGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC 1497
Db |||
Qy 1033 GATGTCGATACCAACGAATTTGCTTAAAAAGGTGAGCAGCTCTTAAACAGCTACGCAACGTAAC 1092
Db |||

RESULT 4
US-09-940-235-1
; Sequence 1, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-09-940-235-1

Query Match      69.8%; Score 1243.4; DB 3; Length 1245;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCATCTGTCAACACAGCCAAATGGTGTGTT 597
Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCATCTGTCAACACAGCCAAATGGTGTGTT 60
Qy 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 657
Db |||
Qy 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 120
Db |||
Qy 658 GATCTAAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 717
Db |||
Qy 121 GATCTAAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
Db |||
Qy 718 AAACATTTGCTACTGATAGTGGCGGATGTACATAAATCTTGAGAAAGCTGACTACTA 777
Db |||

```

```

181 AAACCATTTGCTACTGATAGTGGCGGATGTACATAACTTGAGAAAGCTGACTTACTA 240
778 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTACAGCAGCTACTTTGAGTTC 837
241 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTACAGCAGCTACTTTGAGTTC 300
838 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTACTTTGCTGAC 897
301 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTACTTTGCTGAC 360
898 AAGATGTTTGCAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTACTTTGCTGAC 957
361 AAGATGTTTGCAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTACTTTGCTGAC 420
958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 480
1018 GAATATCTGTACAGTTTACTCCCTTAAACCGCTGATGACGATTTTCAGACGAGTCTCAA 1077
481 GAATATCTGTACAGTTTACTCCCTTAAACCGCTGATGACGATTTTCAGACGAGTCTCAA 540
1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATCTA 1137
541 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATCTA 600
1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
601 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 660
1198 TCCTCAATCGTCACTCATGACAAATGATTTTCGTTACGATTTTACCAATGATCAAGAG 1257
661 TCCTCAATCGTCACTCATGACAAATGATTTTCGTTACGATTTTACCAATGATCAAGAG 720
1258 TTTACTTACCGTTTAAATCGGACACGCTTATAGGATCAATAAATCTGGTCTG 1317
721 TTTACTTACCGTTTAAATCGGACACGCTTATAGGATCAATAAATCTGGTCTG 780
1318 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAGGG 1377
781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATATACGTCCTTAAAAAGGG 840
1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCATCAATACGTT 1437
841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTTGAAACTGTTTACCATCAATACGTT 900
1438 GATGTCGATACCAACGAATGCTTAAAGTGGAGAGCTCTTACAGCTAGGAAAGTAAAC 1497
901 GATGTCGATACCAACGAATGCTTAAAGTGGAGAGCTCTTACAGCTAGGAAAGTAAAC 960
1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAACTCTC 1557
961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACAACTCTC 1020
1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTATAGGAGTAATCAGCATGAC 1617
1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTATAGGAGTAATCAGCATGAC 1080
1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGAGAGAAATGCTAGCTAT 1677
1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGAGAGAAATGCTAGCTAT 1140
1678 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAAACGAGAAATTTACAGTACTCTG 1737
1141 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAAACGAGAAATTTACAGTACTCTG 1200
1738 CGTTATACAGGACACTTATACCTGATTAACCCCTTAAACGACAAATAA 1782
1201 CGTTATACAGGACACTTATACCTGATTAACCCCTTAAACGACAAATAA 1245

```

RESULT 5

US-09-940-235-6

```

; Sequence 6, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-6

```

```

Query Match      69.1%; Score 1232.2; DB 3; Length 1327;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1237; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGCAAGTCTTAAATCTTGAATTC 597
Db 83 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTCTGTAAATAACAGCCNAATGGTGTGT 142
Qy 598 AGCGTGTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATCTTGAATTC 657
Db 143 AGCGTGTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATCTTGAATTC 202
Qy 658 GATCTAATCAATCAACACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 717
Db 203 GATCTAATCAATCAACACCTGCTCATGAGGAAAGACAGACGAGGCTTAAGTCCAAATCA 262
Qy 718 AAACCATTTGCTACTGATGAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 777
Db 263 AAACCATTTGCTACTGATGAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 322
Qy 778 AAGGCTATTCAAGAAACAATTCATCGCTAAACGTCACAGTACAGCAGCTACTTTGAGGTC 837
Db 323 AAGGCTATTCAAGAAACAATTCATCGCTAAACGTCACAGTACAGCAGCTACTTTGAGGTC 382
Qy 838 ATTGATTTTGCAGCGATGCAACCAATTACTGATGATAAGGCTAAACGCGAAGGCTACTTTGCTGAC 897
Db 383 ATTGATTTTGCAGCGATGCAACCAATTACTGATGATAAGGCTAAACGCGAAGGCTACTTTGCTGAC 442
Qy 898 AAGATGTTTGCAGTAACTTTGCCGACCCCAACTGTCGAAAGTATTTTCTAAGCGGACAT 957
Db 443 AAGATGTTTGCAGTAACTTTGCCGACCCCAACTGTCGAAAGTATTTTCTAAGCGGACAT 502
Qy 958 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 503 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 562
Qy 1018 GAATATCTGTACAGTTTACTCCCTTAAACCGCTGATGACGATTTTCAGACGAGTCTCAA 1077
Db 563 GAATATCTGTACAGTTTACTCCCTTAAACCGCTGATGACGATTTTCAGACGAGTCTCAA 622
Qy 1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATCTA 1137
Db 623 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATCTA 682

```



```

QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTGAC 1197
Db 683 GCTCAAGCAAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTGAC 742
QY 1198 TCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db 743 TCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 802
QY 1258 TTTACTTACCGTGTAAATAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 803 TTTACTTACCGTGTAAATAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 862
QY 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db 863 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 922
QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAAAATACGTT 1437
Db 923 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAAAATACGTT 982
QY 1438 GATGCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1497
Db 983 GATGCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAAGCTAGCGAACGTAAC 1042
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1557
Db 1043 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCAACAATCTC 1102
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGGTAGAGGATATACGATGAC 1617
Db 1103 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGGTAGAGGATATACGATGAC 1162
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAAGGATGCTAGCTAT 1677
Db 1163 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCGAAGGATGCTAGCTAT 1222
QY 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGACGAGAGTTTACAGCTACTCG 1737
Db 1223 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGACGAGAGTTTACAGCTACTCG 1282
QY 1738 CGTTATACGGGACACTATACCTGATAACCCCTAACGCAAAATAA 1782
Db 1283 CGTTATACGGGACACTATACCTGATAACCCCTAACGCAAAATAA 1327

```

RESULT 6

```

US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

Query Match      64.5%; Score 1150.2; DB 3; Length 1661;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGTT 597
Db 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCAATTTGGTTGTT 242
QY 598 AGCGTTCCTGCTACTGTTGAGGGGCAAGTAACAGACATTAAGTCTTAATTTTGAATC 657
Db 243 AGCGTTCCTGCTACTGTTGAGGGGCAAGTAACAGACATTAAGTCTTAATTTTGAATC 302
QY 658 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGACGAAAGCTTTAAGTCCAAATCA 717
Db 303 GATCTAACATCAACGACCTGCTCATGGAGGAAAGACAGACGAAAGCTTTAAGTCCAAATCA 362
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTTGAGAAAGCTGACTTACTA 777
Db 363 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTTGAGAAAGCTGACTTACTA 422
QY 778 AAGGCTATTTCAAGNACAAATTTGATCGCTAAAGTCCACAGTAACGAGCTACTTTGAGGTC 837
Db 423 AAGGCTATTTCAAGNACAAATTTGATCGCTAAAGTCCACAGTAACGAGCTACTTTGAGGTC 482
QY 838 ATTGATTTTTCGACGAGTGCACCACTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 897
Db 483 ATTGATTTTTCGACGAGTGCACCACTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 542
QY 898 AAAGATGGTTTGGTAACTTTTCCGACCCCACTGTCCCAAGAAATTTTGTAAAGCGGACAT 957
Db 543 AAAGATGGTTTGGTAACTTTTCCGACCCCACTGTCCCAAGAAATTTTGTAAAGCGGACAT 602
QY 958 GTGCGGTTAGACCATATAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db 603 GTGCGGTTAGACCATATAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTG 662
QY 1018 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCCAGCAGGCTCTCAA 1077
Db 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACCATTTTCCAGCAGGCTCTCAA 722
QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCCGTCGACACCATCATCATCTCAAGAAATTA 1137
Db 723 GATACTAAGCTATTGAAAAACACTAGCTATCCGTCGACACCATCATCATCTCAAGAAATTA 782
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
Db 783 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 842
QY 1198 TCCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1257
Db 843 TCCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 902
QY 1258 TTTTACTTACCGTGTAAATAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 903 TTTTACTTACCGTGTAAATAATCGGGAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962
QY 1318 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db 963 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022
QY 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAAAATACGTT 1437
Db 1023 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATCAAAATACGTT 1082
QY 1438 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAACAGCTACGACGTAAC 1497
Db 1083 GATGTCGATACCAACCAATGCTTAAAGGTGACGAGCTCTTAAACAGCTACGACGTAAC 1142

```


Qy 1498 TTAGACTTCAGAGTTTATAGATCTCTCGTGTAAGGCTAAACTACTCTACAAATCTC 1557
 Db 1143 TTAGACTTCAGAGTTTATAGATCTCTCGTGTAAGGCTAAACTACTCTACAAATCTC 1202
 Qy 1558 GATGCTTTTGGTATTATGACTTAACTTAACTGGAAGCTAGAGGATATCAGATGAC 1617
 Db 1203 GATGCTTTTGGTATTATGACTTAACTTAACTGGAAGCTAGAGGATATCAGATGAC 1262
 Qy 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1677
 Db 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAAATGCTAGCTAT 1322
 Qy 1678 CATTTAGCCCTATGAT 1692
 Db 1323 CATTTAGCCGCTGTT 1337

RESULT 7

US-09-940-235-9
 ; Sequence 9, Application US/09940235
 ; Publication No. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammarra
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 1541
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hybrid cassette
 US-09-940-235-9

Query Match 1 63.6%; Score 1134.2; DB 3; Length 1541;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Qy 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGT 597
 Db 33 ATAGCTGGTCCGTAATGGCTACTAGATCGCTCTGTAAATACAGCCAAATGGTGT 92
 Qy 598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 657
 Db 93 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
 Qy 658 GATCTAAATCATACGACCTGCTCATGAGGAGGAGACAGCAAGGCTTAAAGTCAAAATCA 717
 Db 153 GATCTAAATCATACGACCTGCTCATGAGGAGGAGACAGCAAGGCTTAAAGTCAAAATCA 212
 Qy 718 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAATTCAGAAAGCTGACTACTA 777
 Db 213 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAATTCAGAAAGCTGACTACTA 272
 Qy 778 AAGGCTATTCAAGAAACAAATTTGATCGCTTAACGCTCCACAGTAACGACACTTTTGGGTC 837
 Db 273 AAGGCTATTCAAGAAACAAATTTGATCGCTTAACGCTCCACAGTAACGACACTTTTGGGTC 332

RESULT 8

US-10-474-792-657
 ; Sequence 657, Application US/10474792
 ; Publication No. US20040236072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olmsted, Stephen
 ; APPLICANT: Zagursky, Robert
 ; APPLICANT: Nickbarg, Elliot
 ; APPLICANT: Winter, Lourie
 ; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
 ; FILE REFERENCE: AM 100399

Qy 838 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 897
 Db 333 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 392
 Qy 898 AAAGATGGTTGCGTAAACCTTCCCGACCCAACTGCTGCAAGAAATTTTTCCTAAGCGGACAT 957
 Db 393 AAAGATGGTTGCGTAAACCTTCCCGACCCAACTGCTGCAAGAAATTTTTCCTAAGCGGACAT 452
 Qy 958 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 453 GTGCGCGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 512
 Qy 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGCTCTCAAA 1077
 Db 513 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACAGGCTCTCAAA 572
 Qy 1078 GATATAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 1137
 Db 573 GATATAGCTATTGAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 632
 Qy 1138 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCAGGCTATACGATTTATGAACGTGAC 1197
 Db 633 GCTCAAGCACAAGCATTTTAAACCAAAACCAACCAGGCTATACGATTTATGAACGTGAC 692
 Qy 1198 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257
 Db 693 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
 Qy 1258 TTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
 Db 753 TTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
 Qy 1318 AATGAAGAAATAAACAACACTGATCTCTGAGAAATATACGCTCTTAAAAAAGGG 1377
 Db 813 AATGAAGAAATAAACAACACTGATCTCTGAGAAATATACGCTCTTAAAAAAGGG 872
 Qy 1378 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 1437
 Db 873 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 932
 Qy 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAC 1497
 Db 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGGAAACGTAAC 992
 Qy 1498 TTAGACTTCAGAGATTTTATACGATCTCTGTAAGGCTAAACTACTCTACAAATCTC 1557
 Db 993 TTAGACTTCAGAGATTTTATACGATCTCTGTAAGGCTAAACTACTCTACAAATCTC 1052
 Qy 1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGGATAGAGGATAATCAGATGAC 1617
 Db 1053 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGGATAGAGGATAATCAGATGAC 1112
 Qy 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1677
 Db 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1172
 Qy 1678 CATTTAGCCCTATGAT 1692
 Db 1173 CATTTAGCCGCTGTT 1187

; CURRENT APPLICATION NUMBER: US/10/474,792
 ; CURRENT FILING DATE: 2003-10-14
 ; NUMBER OF SEQ ID NOS: 674
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 657
 ; LENGTH: 1323
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pyogenes
 US-10-474-792-657

Query Match 60.7%; Score 1081.2; DB 8; Length 1323;
 Best Local Similarity 91.7%; Pred. No. 1.3e-299;
 Matches 1143; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy	537	TATTGCTGGACCTGAGTGGCTGTAGACGGCTCATCTGTCAACAAACAGCGCAATGGTTGT	596
Db	78	TATTGCTGGTATGATGGCTACAGACCGTCCACCTATCNAATACAGCCGATGTTGT	137
Qy	597	TAGCGTTGCTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAAT	656
Db	138	TAGTATGGCCGGTATCGTTGAAGGTACCGATAAAAAAGTTTTTATAAATTTTTTTGAAAT	197
Qy	657	CGATCTAACATCAACACCTGCTCAGGAGGAAAGACAGAGCAGGGCTTAAGTCCAAAATC	716
Db	198	CGATCTAACATCAACACCTGCTCAGGAGGAAAGACAGAGCAGGGCTTAAGTCCAAAATC	257
Qy	717	AAAAACATTTGCTACTGATAGTGGCGGATGTACACATAAATTTGAGAAAGCTGACTTACT	776
Db	258	AAAAACATTTGCTACAGATAATGGCGCAATGGCCATAAATTTGAAAGAGCTGACTTATT	317
Qy	777	AAAGCTATTTCGAAGCAATTTGATCGCTTAACCGTCCACAGTAACGACGACTTCTTGAGGT	836
Db	318	AAAGCTATTTCGAAGCAATTTGATCGCTTAACCGTCCACAGTAACGACGACTTCTTGAGGT	377
Qy	837	CATTGATTTGCAAGCAATGCAACCAATTAATGATGCAAGCAAGCGGCTTCTTGCTGA	896
Db	378	CATTGATTTGCAAGCAATGCAACCAATTAATGATGCAAGCAAGCGGCTTCTTGCTGA	437
Qy	897	CAAGATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAGAAATTTTGTGTAAGCGGACA	956
Db	438	CAAGATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAGAAATTTTGTGTAAGCGGACA	497
Qy	957	TGTGCGGCTTAGACCATATAAAGAAAAAACAATACAAAAACCAAGCAAAATCTCTGATGT	1016
Db	498	TGTGCGGCTTAGACCATATAAAGAAAAAACAATACAAAAACCAAGCAAAATCTCTGATGT	557
Qy	1017	GGAATATATGTAAGTTACTTCTTAAACCTGTATGACGATTTTCAAGACCGGCTCAA	1076
Db	558	AGAAATATATGTAAGTTACTTCTTAAACCTGTATGACGATTTTCAAGACCGGCTCAA	617
Qy	1077	AGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA	1136
Db	618	AGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA	677
Qy	1137	AGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCGGCTATACGATTTATCAACGTGA	1196
Db	678	AGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCGGCTATACGATTTATCAACGTGA	737
Qy	1197	CTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTGACGATTTTCAATGATCAAGA	1256
Db	738	CTCCTCAATCGTCACTCATGACAAATGACATTTTCCGTGACGATTTTCAATGATCAAGA	797
Qy	1257	GTTTACTTACCGTGTAAATAATCGGGAAACAGCTTATAGATCAATTAATAAATCTCGGTCT	1316
Db	798	GTTTACTTACCGTGTAAATAATCGGGAAACAGCTTATAGATCAATTAATAAATCTCGGTCT	857
Qy	1317	GAAACAAGAAATAAACAACCTGATCTCTGAGAAATATATAGTCTCTTAAAGAGG	1376
Db	858	TAAAGAAAAACGAACAACACTGATCTGGTCTCTGAGAAATATATAGTCTCTTAAAGAGG	917
Qy	1377	GGAAAGCGGTATGATCCCTTTTGTGCGAGTCACTTTGAAACTGTTCCACCATCAATACGT	1436
Db	918	GGAAAGCGGTATGATCCCTTTTGTGCGAGTCACTTTGAAACTGTTCCACCATCAATACGT	977

RESULT 9

US-10-210-120-49
 ; Sequence 49, Application US/10210120
 ; Publication No. US20030175736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chinnaiyan, Arul M.
 ; APPLICANT: Rubin, Mark A.
 ; APPLICANT: Sreekumar, Arun
 ; TITLE OF INVENTION: Expression Profile of Prostate Cancer
 ; FILE REFERENCE: UM-07221
 ; CURRENT APPLICATION NUMBER: US/10/210,120
 ; CURRENT FILING DATE: 2002-08-01
 ; PRIOR APPLICATION NUMBER: US 60/309,581
 ; PRIOR FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: US 60/334,468
 ; PRIOR FILING DATE: 2001-11-15
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 49
 ; LENGTH: 2127
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-210-120-49

Query Match 13.1%; Score 232.8; DB 6; Length 2127;
 Best Local Similarity 79.8%; Pred. No. 2e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

Qy	254	CCCATAGCTCAGAAAGTCTTTTGGATCATGCTGCGGACTTCTCTATGTTGGTGGGAAACG	313
Db	597	CCCATAGCTCAGAAAGTCTTTTGGATCATGCTGCGGACTTCTCTATGTTGGTGGGAAACG	656
Qy	314	-----TGGGAGAAAGGCAGC	327
Db	657	TGGGAGAAAGCCCTACCAAGGCTGGATGTTGTACTTCCCTGGGAGAGGCAGC	716
Qy	328	GGACGCATCATTTGCACTTTAGAAATAGATGCAACGATCAGGACACAAGGACATCTTAT	387
Db	717	GGACGCATCATTTGCACTTTAGAAATAGATGCAACGATCAGGACACAAGGACATCTTAT	776
Qy	388	AGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGCTCCAGTGCATCTGC	447
Db	777	AGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGCTCCAGTGCATCTGC	836
Qy	448	ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCG	507
Db	837	ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGACACCTCTGTGAGACCAACATCG	896

QY 508 AGCGATCTGGCCCTTACCGATGTTCTGTTATCTGAGCCTGAGTGGCTGCTAGACCGT 567
 Db |||||
 QY 897 AGCGATCTGGCCCTTACCGATGTTCTGTTATCTGAGCCTGAGTGGCTGCTAGACCGT 566
 Db |||||
 QY 568 CCATCTGTCAACAACAGCCCAATTGGT 593
 Db |||||
 QY 957 CAGCCTCTCCCTATGCGCCACTGTGT 982
 Db |||||

RESULT 10
 US-10-956-157-4288
 ; Sequence 4288, Application US/10956157
 ; Publication No. US20050118625A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
 ; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
 ; CURRENT FILING DATE: 2004-10-04
 ; NUMBER OF SEQ ID NOS: 319805
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4288
 ; LENGTH: 2127
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-956-157-4288

Query Match 13.1%; Score 232.8; DB 9; Length 2127;
 Best Local Similarity 79.8%; Pred. No. 2e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
 QY 254 CCCATAGCTGAGAGTGTCTTTCATCATCTGCTGGGACTTCTATGTTGGTGGGAGAAACG 313
 Db |||||
 QY 597 CCCATAGCTGAGAGTGTCTTTCATCATCTGCTGGGACTTCTATGTTGGTGGGAGAAACG 656
 Db |||||
 QY 314 -----TGGGAGAGGCGAGC 327
 Db |||||
 QY 657 TGGGAGAGGCGCTACCAAGCTGGATGATGTTAGTGTCTGCTGGGAGAGGCGAGC 716
 Db |||||
 QY 328 GGACGATCATCTTGCATCTTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTAT 387
 Db |||||
 QY 717 GGACGATCATCTTGCATCTTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTAT 776
 Db |||||
 QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGAAACCTCTCCAGTGCATCTGC 447
 Db |||||
 QY 777 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGAAACCTCTCCAGTGCATCTGC 836
 Db |||||
 QY 448 ACAGGCAACGCGGAGGAGAGTGGAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 507
 Db |||||
 QY 837 ACAGGCAACGCGGAGGAGAGTGGAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 896
 Db |||||
 QY 508 AGCGATCTGGCCCTTACCGATGTTCTGTTATCTGAGCCTGAGTGGCTGCTAGACCGT 567
 Db |||||
 QY 897 AGCGATCTGGCCCTTACCGATGTTCTGTTATCTGAGCCTGAGTGGCTGCTAGACCGT 567
 Db |||||
 QY 568 CCATCTGTCAACAACAGCCCAATTGGT 593
 Db |||||
 QY 957 CAGCCTCTCCCTATGCGCCACTGTGT 982
 Db |||||

RESULT 11
 US-10-909-035-49
 ; Sequence 49, Application US/10909035
 ; Publication No. US20050136493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rubin, Mark A.
 ; APPLICANT: Chinnaiyan, Arul M.
 ; APPLICANT: Laxman, Bharathi
 ; APPLICANT: Sreekumar, Arun
 ; TITLE OF INVENTION: AMACR Cancer Markers

FILE REFERENCE: UM-09098
 ; CURRENT APPLICATION NUMBER: US/10/909,035
 ; CURRENT FILING DATE: 2004-07-30
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 49
 ; LENGTH: 2127
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-909-035-49

Query Match 13.1%; Score 232.8; DB 9; Length 2127;
 Best Local Similarity 79.8%; Pred. No. 2e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
 QY 254 CCCATAGCTGAGAGTGTCTTTCATCATCTGCTGGGACTTCTATGTTGGTGGGAGAAACG 313
 Db |||||
 QY 597 CCCATAGCTGAGAGTGTCTTTCATCATCTGCTGGGACTTCTATGTTGGTGGGAGAAACG 656
 Db |||||
 QY 314 -----TGGGAGAGGCGAGC 327
 Db |||||
 QY 657 TGGGAGAGGCGCTACCAAGCTGGATGATGTTAGTGTCTGCTGGGAGAGGCGAGC 716
 Db |||||
 QY 328 GGACGATCATCTTGCATCTTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTAT 387
 Db |||||
 QY 717 GGACGATCATCTTGCATCTTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTAT 776
 Db |||||
 QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGAAACCTCTCCAGTGCATCTGC 447
 Db |||||
 QY 777 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGAAACCTCTCCAGTGCATCTGC 836
 Db |||||
 QY 448 ACAGGCAACGCGGAGGAGAGTGGAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 507
 Db |||||
 QY 837 ACAGGCAACGCGGAGGAGAGTGGAGTGTGAGAGGACACCTCTGTGAGACCAACATCG 896
 Db |||||
 QY 508 AGCGATCTGGCCCTTACCGATGTTCTGTTATCTGAGCCTGAGTGGCTGCTAGACCGT 567
 Db |||||
 QY 897 AGCGATCTGGCCCTTACCGATGTTCTGTTATCTGAGCCTGAGTGGCTGCTAGACCGT 567
 Db |||||
 QY 568 CCATCTGTCAACAACAGCCCAATTGGT 593
 Db |||||
 QY 957 CAGCCTCTCCCTATGCGCCACTGTGT 982
 Db |||||

RESULT 12
 US-10-741-601-70
 ; Sequence 70, Application US/10741601
 ; Publication No. US20040166519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 26415
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 70
 ; LENGTH: 2443
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-741-601-70

Query Match 13.1%; Score 232.8; DB 7; Length 2443;
 Best Local Similarity 79.8%; Pred. No. 2.1e-55;
 Matches 308; Conservative 0; Mismatches 33; Indels 46; Gaps 1;
 QY 254 CCCATAGCTGAGAGTGTCTTTCATCATCTGCTGGGACTTCTATGTTGGTGGGAGAAACG 313
 Db |||||
 QY 913 CCCATAGCTGAGAGTGTCTTTCATCATCTGCTGGGACTTCTATGTTGGTGGGAGAAACG 972
 Db |||||
 QY 314 -----TGGGAGAGGCGAGC 327
 Db |||||

Db 973 TGGGAGAACCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCAGC 1032
 Qy 328 GGAAGCATCACTTGCCTCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 387
 Db 1033 GGAAGCATCACTTGCCTCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 1092
 Qy 388 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 447
 Db 1093 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
 Qy 448 ACAGGCAACGGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGAGACACATCG 507
 Db 1153 ACAGGCAACGGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGAGACACATCG 1212
 Qy 508 AGCGGATCTGGCCCTTACCGATGTTCTGTTGCTGGACCTGAGTGGCTGTAGACCGT 567
 Db 1213 AGCGGATCTGGCCCTTACCGATGTTCTGTTGCTGGACCTGAGTGGCTGTAGACCGT 1272
 Qy 568 CCATCTGTCAACAAACAGCCAAATTGGT 593
 Db 1273 CAGCCTCTCCCTATGCGCACTGTGT 1298

RESULT 13
 US-10-741-600-238
 ; Sequence 238, Application US/10741600
 ; Publication No. US20050026169A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CL001499
 ; CURRENT APPLICATION NUMBER: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 73997
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 238
 ; LENGTH: 2443
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-741-600-238

Query Match 13.1%; Score 232.8; DB 8; Length 2443;
 Best Local Similarity 79.8%; Pred. No. 2.1e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
 Qy 254 CCCATAGCTGAGAAGTGTCTTGTATCATGCTGCTGGGACTTCTCTATGTTGGTGGGAGAAACG 313
 Db 913 CCCATAGCTGAGAAGTGTCTTGTATCATGCTGCTGGGACTTCTCTATGTTGGTGGGAGAAACG 972
 Qy 314 -----TGGGAGAGGCAGC 327
 Db 973 TGGGAGAACCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCAGC 1032
 Qy 328 GGAAGCATCACTTGCCTCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 387
 Db 1033 GGAAGCATCACTTGCCTCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 1092
 Qy 388 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 447
 Db 1093 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
 Qy 448 ACAGGCAACGGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGAGACACATCG 507
 Db 1153 ACAGGCAACGGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGAGACACATCG 1212
 Qy 508 AGCGGATCTGGCCCTTACCGATGTTCTGTTGCTGGACCTGAGTGGCTGTAGACCGT 567
 Db 1213 AGCGGATCTGGCCCTTACCGATGTTCTGTTGCTGGACCTGAGTGGCTGTAGACCGT 1272
 Qy 568 CCATCTGTCAACAAACAGCCAAATTGGT 593
 Db 1273 CAGCCTCTCCCTATGCGCACTGTGT 1298

RESULT 14
 US-10-741-601-75
 ; Sequence 75, Application US/10741601
 ; Publication No. US20040166519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT APPLICATION NUMBER: US/10741,601
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 26415
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 75
 ; LENGTH: 2488
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-741-601-75

Query Match 13.1%; Score 232.8; DB 7; Length 2488;
 Best Local Similarity 79.8%; Pred. No. 2.1e-55;
 Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
 Qy 254 CCCATAGCTGAGAAGTGTCTTGTATCATGCTGCTGGGACTTCTCTATGTTGGTGGGAGAAACG 313
 Db 913 CCCATAGCTGAGAAGTGTCTTGTATCATGCTGCTGGGACTTCTCTATGTTGGTGGGAGAAACG 972
 Qy 314 -----TGGGAGAGGCAGC 327
 Db 973 TGGGAGAACCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCAGC 1032
 Qy 328 GGAAGCATCACTTGCCTCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 387
 Db 1033 GGAAGCATCACTTGCCTCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT 1092
 Qy 388 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 447
 Db 1093 AGAATTGGAGACACCTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
 Qy 448 ACAGGCAACGGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGAGACACATCG 507
 Db 1153 ACAGGCAACGGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGAGACACATCG 1212
 Qy 508 AGCGGATCTGGCCCTTACCGATGTTCTGTTGCTGGACCTGAGTGGCTGTAGACCGT 567
 Db 1213 AGCGGATCTGGCCCTTACCGATGTTCTGTTGCTGGACCTGAGTGGCTGTAGACCGT 1272
 Qy 568 CCATCTGTCAACAAACAGCCAAATTGGT 593
 Db 1273 CAGCCTCTCCCTATGCGCACTGTGT 1298

RESULT 15
 US-10-741-600-244
 ; Sequence 244, Application US/10741600
 ; Publication No. US20050026169A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT APPLICATION NUMBER: US/10741,600
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 73997
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 244
 ; LENGTH: 2488
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-741-600-244

```

Query Match      13.1%; Score 232.8; DB 8; Length 2488;
Best Local Similarity 79.8%; Pred. No. 2.1e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

QY 254 CCATAGCTGAGAGTGTGTTGATCATGCTGGGACTTCCTATGTGTCGGAGAAACG 313
Db 913 CCATAGCTGAGAGTGTGTTGATCATGCTGGGACTTCCTATGTGTCGGAGAAACG 972
QY 314 -----TGGGAGAAAGGCAGC 327
Db 973 TGGGAGAAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCAGC 1032
QY 328 GGACGCATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 387
Db 1033 GGACGCATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCCAGTGCATCTGC 447
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCCAGTGCATCTGC 1152
QY 448 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG 507
Db 1153 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG 1212
QY 508 AGCGATCTGGCCCTTTCACCGATGTTTCGATTGCTGGACCTGAGTGGCTGTAGACCGT 567
Db 1213 AGCGATCTGGCCCTTTCACCGATGTTTCGATTGCTGGACCTGAGTGGCTGTAGACCGT 1272
QY 568 CCATCTGTCAACACAGCCAATTGGT 593
Db 1273 CAGCCTCCTCCTTATGGCCACTGTGT 1298

```

Search completed: February 1, 2006, 14:21:21
Job time : 1474.42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 18:17:59 ; Search time 329.469 Seconds
(without alignments)
4492.841 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tcgcttcacgttcgctcgcg.....ataacctaacgacaaataa 1782

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6059551 seqs, 415333918 residues

Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1782	100.0	1782	US-10-631-558-11	Sequence 11, Appl
2	1694	94.5	2096	US-10-631-558-12	Sequence 12, Appl
3	1245	69.9	1377	US-10-631-558-5	Sequence 5, Appl
4	1243.4	69.8	1245	US-10-631-558-1	Sequence 1, Appl
5	1232.2	69.1	1327	US-10-631-558-6	Sequence 6, Appl
6	1150.2	64.5	1661	US-10-631-558-10	Sequence 10, Appl
7	1134.2	63.6	1541	US-10-631-558-9	Sequence 9, Appl
8	232.8	13.1	2443	US-10-995-561-114	Sequence 114, App
9	232.8	13.1	2488	US-10-995-561-105	Sequence 105, App
10	232.8	13.1	6510	US-10-995-561-112	Sequence 112, App
11	232.8	13.1	7823	US-10-995-561-117	Sequence 117, App
12	232.8	13.1	7848	US-10-995-561-111	Sequence 111, App
13	232.8	13.1	7935	US-10-995-561-113	Sequence 113, App
14	232.8	13.1	7959	US-10-995-561-108	Sequence 108, App
15	232.8	13.1	8013	US-10-995-561-104	Sequence 104, App
16	232.8	13.1	8155	US-10-995-561-116	Sequence 116, App
17	232.8	13.1	8226	US-10-995-561-107	Sequence 107, App
18	232.8	13.1	8232	US-10-821-234-693	Sequence 693, App
19	232.8	13.1	8278	US-10-995-561-106	Sequence 106, App
20	232.8	13.1	8332	US-10-995-561-110	Sequence 110, App
21	232.8	13.1	8371	US-10-995-561-109	Sequence 109, App
22	228	12.8	777	US-10-631-558-3	Sequence 3, Appl

23	160.8	9.0	87672	7	US-10-995-561-13237	Sequence 13237, A
24	142	8.0	8404	8	US-11-136-527-2446	Sequence 2446, Ap
c 25	134	7.5	3122	8	US-11-009-840A-416	Sequence 416, App
c 26	134	7.5	3122	8	US-11-009-873A-416	Sequence 416, App
c 27	134	7.5	3618	8	US-11-009-840A-72	Sequence 72, Appl
c 28	134	7.5	3618	8	US-11-009-873A-72	Sequence 72, Appl
c 29	134	7.5	4458	7	US-10-742-634-15	Sequence 15, Appl
c 30	134	7.5	4980	7	US-10-742-634-15	Sequence 15, Appl
c 31	134	7.5	5678	8	US-11-106-820-17	Sequence 17, Appl
c 32	134	7.5	5678	8	US-11-106-820-18	Sequence 18, Appl
c 33	134	7.5	5679	8	US-11-106-820-13	Sequence 13, Appl
c 34	134	7.5	5679	8	US-11-106-820-14	Sequence 14, Appl
c 35	134	7.5	12391	8	US-11-017-550-67	Sequence 67, Appl
c 36	134	7.5	16360	7	US-10-519-531-1	Sequence 1, Appl
c 37	134	7.5	17207	7	US-10-519-531-8	Sequence 8, Appl
c 38	132.4	7.4	11627	8	US-11-230-995-1	Sequence 1, Appl
c 39	131	7.4	9808	7	US-10-966-483-43	Sequence 43, Appl
c 40	131	7.4	9808	8	US-11-021-441-33	Sequence 33, Appl
c 41	110.6	6.2	5347	8	US-11-209-589-1	Sequence 1, Appl
c 42	110.6	6.2	8031	7	US-10-623-155-254	Sequence 254, App
c 43	100.4	5.6	4974	8	US-11-094-586-17	Sequence 17, Appl
c 44	100.4	5.6	4974	8	US-11-076-733-86	Sequence 86, Appl
c 45	100.4	5.6	8966	8	US-11-076-733-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-10-631-558-11
; Sequence 11, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631.558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-11

Query Match 100.0%; Score 1782; DB 7; Length 1782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCGCTTCACGTTTCGCTCGGTATCGGTGATTCATTCTGCTAACAGTAAGCAACCCGCG 60
Db 1 TCGCTTCACGTTTCGCTCGGTATCGGTGATTCATTCTGCTAACAGTAAGCAACCCGCG 60
Qy 61 CAGCCTAGCCGGTCTCTCAACGACGAGACGATCATGCCACCCGTCGCGACGACCA 120

Db 61 CAGCCCTAGCCGGTCTCTCAACGACAGGACGATCATGCGCACCCCGTGGCCAGGACCCA 120
Qy 121 AGCCTGCCGAGATCTCGATCCGCGAAATTAATACGACTCACTATAGGGAGACCAAC 180
Db 121 AGCCTGCCGAGATCTCGATCCGCGAAATTAATACGACTCACTATAGGGAGACCAAC 180
Qy 181 GGTTCCTCTAGAAATAATTTTGTAACTTTAAGAAAGGAGATATACCATGTCGAAGC 240
Db 181 GGTTCCTCTAGAAATAATTTTGTAACTTTAAGAAAGGAGATATACCATGTCGAAGC 240
Qy 241 ACAACAGATGTACCATAGCTGAGAAGTGTGTTGATCATGCTGGGACTTCTCTATGT 300
Db 241 ACAACAGATGTACCATAGCTGAGAAGTGTGTTGATCATGCTGGGACTTCTCTATGT 300
Qy 301 GGTTCGAGAAACGTCGGAGAGGACGCGGACGCGATCACTTGGCATCTTAGAATAGATGC 360
Db 301 GGTTCGAGAAACGTCGGAGAGGACGCGGACGCGATCACTTGGCATCTTAGAATAGATGC 360
Qy 361 AACGATCAGGACACAAAGGACATCCTATAGAAATGGAGACACCTGGAGCAAGAGGATAAT 420
Db 361 AACGATCAGGACACAAAGGACATCCTATAGAAATGGAGACACCTGGAGCAAGAGGATAAT 420
Qy 421 CGAGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGGAGAGTGGAAGTGTGAG 480
Db 421 CGAGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGGAGAGTGGAAGTGTGAG 480
Qy 481 AGGCACACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTTCAACCGATGTCGTATT 540
Db 481 AGGCACACCTCTGTGCGAGACCAATCGAGCGGATCTGGCCCTTCAACCGATGTCGTATT 540
Qy 541 GCTGACCTGAGTGGCTGTAGACCTTCCATCTGTCAACAGCAACCAATTTGGTTGTAGC 600
Db 541 GCTGACCTGAGTGGCTGTAGACCTTCCATCTGTCAACAGCAACCAATTTGGTTGTAGC 600
Qy 601 GTTGCTGTGTAATGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAAATCGAT 660
Db 601 GTTGCTGTGTAATGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGGAAATCGAT 660
Qy 661 CTAAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAA 720
Db 661 CTAAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAA 720
Qy 721 CCATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTACTTAAAG 780
Db 721 CCATTTGCTACTGATAGTGGCGGATGTCATATAAATTTGAGAAAGCTACTTAAAG 780
Qy 781 GCTATTCAAGAAACATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGGTCATT 840
Db 781 GCTATTCAAGAAACATTTGATCGCTAACGTCACAGTAAACGACGACTACTTTGAGGTCATT 840
Qy 841 GATTTTGAAGCGATGCAACCATTAATGATCGAAACGCGCAAGGCTACTTTGCTGCAAAA 900
Db 841 GATTTTGAAGCGATGCAACCATTAATGATCGAAACGCGCAAGGCTACTTTGCTGCAAAA 900
Qy 901 GATGTTGCGTAACTTGGCGACCCCACTGTCGAAGATTTTGTGTAAGCGGACATGTG 960
Db 901 GATGTTGCGTAACTTGGCGACCCCACTGTCGAAGATTTTGTGTAAGCGGACATGTG 960
Qy 961 CCGGTTAGACCATATAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTGAA 1020
Db 961 CCGGTTAGACCATATAAGAAACCAATATACAAACCAAGCGAAATCTGTTGATGTGAA 1020
Qy 1021 TATACGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTTCAAGAGTCTCAAGAT 1080
Db 1021 TATACGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTTCAAGAGTCTCAAGAT 1080
Qy 1081 ACTAGCTATTGAAACATAGCTATCGGTGACACCATCATCTCAAGATTTACTAGCT 1140
Db 1081 ACTAGCTATTGAAACATAGCTATCGGTGACACCATCATCTCAAGATTTACTAGCT 1140
Qy 1141 CAAGCAAAAGCATTTTAAACCAAAACCAAGGCTATACGATTTATGAACTGACTCC 1200
Db 1141 CAAGCAAAAGCATTTTAAACCAAAACCAAGGCTATACGATTTATGAACTGACTCC 1200

Qy 1201 TCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGCGATCAAGAGTTT 1260
Db 1201 TCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGCGATCAAGAGTTT 1260
Qy 1261 ACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTGAAT 1320
Db 1261 ACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTGAAT 1320
Qy 1321 GAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGGGAA 1380
Db 1321 GAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGGGAA 1380
Qy 1381 AAGCGTATCATCCCTTTGATCCGAGTCACCTTGAACCTGTTCCACATCAATAGCTTGAT 1440
Db 1381 AAGCGTATCATCCCTTTGATCCGAGTCACCTTGAACCTGTTCCACATCAATAGCTTGAT 1440
Qy 1441 GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTA 1500
Db 1441 GTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTTA 1500
Qy 1501 GACTTCAGAGATTTTATACGATCTCTGTAAGGCTAAACTTACTCTCAACAATCTCGAT 1560
Db 1501 GACTTCAGAGATTTTATACGATCTCTGTAAGGCTAAACTTACTCTCAACAATCTCGAT 1560
Qy 1561 GCTTTTGGTATTATGGACTATACCTTAACTGGGAAAGTAGAGGATAATCAGATGACACC 1620
Db 1561 GCTTTTGGTATTATGGACTATACCTTAACTGGGAAAGTAGAGGATAATCAGATGACACC 1620
Qy 1621 AACCGTATCATAACCGTTTATATGGCAAGCGACCGAAGGAGAGAATGCTAGCTATCAT 1680
Db 1621 AACCGTATCATAACCGTTTATATGGCAAGCGACCGAAGGAGAGAATGCTAGCTATCAT 1680
Qy 1681 TTAGCTTATGATAAAGATCGTTTATACCGAAGAAGACGAGAAGTTTACGCTACCTCGGT 1740
Db 1681 TTAGCTTATGATAAAGATCGTTTATACCGAAGAAGACGAGAAGTTTACGCTACCTCGGT 1740
Qy 1741 TATACGGGACACCTTATACCTGATAACCTTAAACGACAAATAA 1782
Db 1741 TATACGGGACACCTTATACCTGATAACCTTAAACGACAAATAA 1782

RESULT 2

US-10-631-558-12
; Sequence 12, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence


```

; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-12

```

Query Match 94.5%; Score 1684; DB 7; Length 2096;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1	TCGCTT	CA	CGTTT	CGCT	CGCT	AT	CGGT	GA	TAT	CT	CT	GA	TA	CG	TA	AG	CA	AC	CG	CG	60
51	TCGCTT	CA	CGTTT	CGCT	CGCT	AT	CGGT	GA	TAT	CT	CT	GA	TA	CG	TA	AG	CA	AC	CG	CG	110
61	CAGCTT	AG	CGGGT	CT	CA	CG	AC	GAG	AG	CA	CG	AT	CA	TG	CG	CA	CC	CG	TG	GC	120
111	CAGCTT	AG	CGGGT	CT	CA	CG	AC	GAG	AG	CA	CG	AT	CA	TG	CG	CA	CC	CG	TG	GC	170
121	ACGTTG	CCCG	AGAT	CT	CG	AT	CCCG	CG	AAAT	TT	TA	AT	CA	CG	AT	CT	AT	AT	AG	GAG	180
171	ACGTTG	CCCG	AGAT	CT	CG	AT	CCCG	CG	AAAT	TT	TA	AT	CA	CG	AT	CT	AT	AT	AG	GAG	230
181	GGTTT	CCCT	CT	AG	AAAT	TA	TTT	TG	TT	TA	CT	TT	AA	GA	AG	AG	AT	AT	AC	TAT	240
231	GGTTT	CCCT	CT	AG	AAAT	TA	TTT	TG	TT	TA	CT	TT	AA	GA	AG	AG	AT	AT	AC	TAT	290
241	ACAA	CAG	ATT	TG	TAC	CC	TA	GT	CT	G	A	A	AG	T	G	T	T	T	G	A	300
291	ACAA	CAG	ATT	TG	TAC	CC	TA	GT	CT	G	A	A	AG	T	G	T	T	T	G	A	350
301	GGT	CG	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	360
351	GGT	CG	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	410
361	A	A	C	G	A	T	C	A	G	A	C	A	T	C	T	A	T	A	G	A	420
411	A	A	C	G	A	T	C	A	G	A	C	A	T	C	T	A	T	A	G	A	470
421	C	G	A	G	A	A	A	C	T	G	C	T	C	A	G	G	C	A	A	C	480
471	C	G	A	G	A	A	C	T	G	C	T	C	A	G	G	C	A	A	C	G	530
481	A	G	G	C	A	C	T	T	G	T	G	C	A	G	A	C	A	T	C	G	540
531	A	G	G	C	A	C	T	T	G	T	G	C	A	G	A	C	A	T	C	G	590
541	G	C	T	T	G	A	C	C	T	G	A										

[illegible]

RESULT 3
 US-10-631-558-5
 ; Sequence 5, Application US/10631558
 ; Publication No. US20050260598A1
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ;
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ;
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ;
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ;
 ; TITLE OF INVENTION: PROTEIN
 ;
 ; FILE REFERENCE: 07064-009002
 ;
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ;
 ; CURRENT FILING DATE: 2003-07-31
 ;
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ;
 ; PRIOR FILING DATE: 2002-04-09

```

; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-631-558-5

Query Match          69.9%; Score 1245; DB 7; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACAGCCAAATGGTGTGTT 597
DB 133 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACAGCCAAATGGTGTGTT 192

QY 598 AGCGTTGCTGCTACTCTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 657
DB 193 AGCGTTGCTGCTACTCTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 252

QY 658 GATCTAAACATCACGACCTGCTCATGGAGGAAACAGACAGCAAGCGCTTAAAGTCCAAATCA 717
DB 253 GATCTAAACATCACGACCTGCTCATGGAGGAAACAGACAGCAAGCGCTTAAAGTCCAAATCA 312

QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 777
DB 313 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAAACTTGAGAAAGCTGACTTACTA 372

QY 778 AAGGCTATTCAAGAAACAATGATCGCTAAGCTCCACAGTAACGACGACTACTTTGAGTGC 837
DB 373 AAGGCTATTCAAGAAACAATGATCGCTAAGCTCCACAGTAACGACGACTACTTTGAGTGC 432

QY 838 ATTGATTTTGAAGCGCATGCAACCATTTACTGTATCGAAACGGCAAGGTCTACTTTGCTGAC 897
DB 433 ATTGATTTTGAAGCGCATGCAACCATTTACTGTATCGAAACGGCAAGGTCTACTTTGCTGAC 492

QY 898 AAGATGTTTGGTAACTTTGCCGACCAACCTGTGCCAAGATTTTGTCTAAGCGGACAT 957
DB 493 AAGATGTTTGGTAACTTTGCCGACCAACCTGTGCCAAGATTTTGTCTAAGCGGACAT 552

QY 958 GTGCGGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG 1017
DB 553 GTGCGGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGTGATGTG 612

QY 1018 GAATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGTCTCAAA 1077
DB 613 GAATACGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGTCTCAAA 672

QY 1078 GATACTAAGCTATTGAAACACATAGCTATCCGTTGACACCATCATCTCAAGAAATTA 1137
DB 673 GATACTAAGCTATTGAAACACATAGCTATCCGTTGACACCATCATCTCAAGAAATTA 732

QY 1138 GCTCAAGCAACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
DB 733 GCTCAAGCAACAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTGAC 792

QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCATGGATCAAGAG 1257
DB 793 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCATGGATCAAGAG 852

QY 1258 TTTACTTACCGTGTGTTAAATAACGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
DB 853 TTTACTTACCGTGTGTTAAATAACGGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 912

QY 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 1377
DB 913 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGGG 972

QY 1378 GAAAAGCCGTATGATCCCTTTTGTATCGAGTCATCTGAAACTGTTTACCATCAAAATAGCTT 1437

```

```

RESULT 4
US-10-631-558-1
; Sequence 1, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631.558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-10-631-558-1

```

```

Query Match          69.8%; Score 1243.4; DB 7; Length 1245;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACAGCCAAATGGTGTGTT 597
DB 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACAGCCAAATGGTGTGTT 60

QY 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAATC 657

```

Db 61 AGCGTTGCTGGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATATTTTGGAAATC 120
 Qy 658 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 717
 Db 121 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 180
 Qy 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTTACTA 777
 Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTTACTA 240
 Qy 778 AAGGCTATTCAAGAAATGATGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 837
 Db 241 AAGGCTATTCAAGAAATGATGCTTAAGTCCACAGTAACGACGACTACTTTGAGGTC 300
 Qy 838 ATTGATTTTGGCAAGCATGCAACATTAAGTCCGAAACCGCAAGGCTTACTTTGCTGAC 897
 Db 301 ATTGATTTTGGCAAGCATGCAACATTAAGTCCGAAACCGCAAGGCTTACTTTGCTGAC 360
 Qy 898 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 957
 Db 361 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 420
 Qy 958 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
 Db 421 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 480
 Qy 1018 GAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACCATTTTACAGCAGGCTCAAA 1077
 Db 481 GAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACCATTTTACAGCAGGCTCAAA 540
 Qy 1078 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 1137
 Db 541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 600
 Qy 1138 GCTCAAGCAAAAGCAATTTAAACAAACCAAGGCTATACGATTTATGAAACGTCAC 1197
 Db 601 GCTCAAGCAAAAGCAATTTAAACAAACCAAGGCTATACGATTTATGAAACGTCAC 660
 Qy 1198 TCCTCAATCGTCACTACATGACATGACATTTCCGTACGATTTTACCAATGATCAAG 1257
 Db 661 TCCTCAATCGTCACTACATGACATGACATTTCCGTACGATTTTACCAATGATCAAG 720
 Qy 1258 TTTTACTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAATCTGGTCTG 1317
 Db 721 TTTTACTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAATCTGGTCTG 780
 Qy 1318 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1377
 Db 781 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 840
 Qy 1378 GAAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAATCTTTACCATCAATAGCTT 1437
 Db 841 GAAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAATCTTTACCATCAATAGCTT 900
 Qy 1438 GATGTCGATACCAACGAATTTGATAAGTGAAGTACGAGCTTAAACAGTACGAGCACTAAC 1497
 Db 901 GATGTCGATACCAACGAATTTGATAAGTGAAGTACGAGCTTAAACAGTACGAGCACTAAC 960
 Qy 1498 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTAAATCTCTACAAATCTC 1557
 Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATGAGGCTAAATCTCTACAAATCTC 1020
 Qy 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATAATACGATGAC 1617
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGTAGAGGATAATACGATGAC 1080
 Qy 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGGAAGAGAGAAATGCTAGCTAT 1677
 Db 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGGAAGAGAGAAATGCTAGCTAT 1140
 Qy 1678 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAACGAGAAATTTACGCTACTG 1737

Db 1141 CATTTAGCTTATGATAAAGATCGTTATATACCGAAGAAACGAGAAATTTACAGCTACCTG 1200
 Qy 1738 CGTTATACAGGACACCTTATACCTGATACCTTAACCTTAACGACAAATAA 1782
 Db 1201 CGTTATACAGGACACCTTATACCTGATACCTTAACCTTAACGACAAATAA 1245

RESULT 5
 US-10-631-558-6
 ; Sequence 6, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631.558
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940.235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1327
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; US-10-631-558-6

Query Match 69.1%; Score 1232.2; DB 7; Length 1327;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1237; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 538 ATTGCTGGACTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTTGGTTGT 597
 Db 83 ATAGCTGGTCTGTAATGGCTACTAGATCGTCTTCTGTAAATAACAGCAATTTGGTTGT 142
 Qy 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 657
 Db 143 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 202
 Qy 658 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGACGAAGGCTTAAGTCCAAATCA 717
 Db 203 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGACGAAGGCTTAAGTCCAAATCA 262
 Qy 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTTACTA 777
 Db 263 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAAATTTGAGAAAGCTGACTTACTA 322
 Qy 778 AAGGCTATTCAAGAAATTTGATCGCTTAAACCGTCAACAGTACGACGACTTCTTTGAGGTC 837
 Db 323 AAGGCTATTCAAGAAATTTGATCGCTTAAACCGTCAACAGTACGACGACTTCTTTGAGGTC 382
 Qy 838 ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 897
 Db 383 ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 442
 Qy 898 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 957
 Db 443 AAGAGTGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTAAAGCGGACAT 502
 Qy 958 GTGCGGTTAGACCATATAAGAAACCAATACAAACCAAGCAAAATCTGTTGATGTG 1017

Db 503 GTGCGGTTAGACCATATAAGAAAAAACAATACAAAACCAAGCGAATCTGTTGATGTG 562
Qy 1018 GAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 1077
Db 563 GAATATACGTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 622
Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCCGTTGACACCATCATCATCTCAAGAAATTA 1137
Db 623 GATACTAAGCTATTGAAAAACACTAGCTATCCGTTGACACCATCATCATCTCAAGAAATTA 682
Qy 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTTGAC 1197
Db 683 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTTGAC 742
Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 1257
Db 743 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 802
Qy 1258 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 803 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 862
Qy 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1377
Db 863 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 922
Qy 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCATCAATACGTT 1437
Db 923 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTACCATCAATACGTT 982
Qy 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 1497
Db 983 GATGTCGATACCAACGAATTTGCTAAAAAGTGACGAGCTCTTAACAGCTAGCGAACGTAAC 1042
Qy 1498 TTAGACTTCAGAGATTTATACGATCTCTGTTAAGCTAAACCTCTACACAAATCTC 1557
Db 1043 TTAGACTTCAGAGATTTATACGATCTCTGTTAAGCTAAACCTCTACACAAATCTC 1102
Qy 1558 GATGCTTTTGGTATTTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1617
Db 1103 GATGCTTTTGGTATTTATGACATATACCTTAACTGGAAAAAGTAGAGGATATACGATGAC 1162
Qy 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACGAGGAGATGCTAGCTAT 1677
Db 1163 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACGAGGAGATGCTAGCTAT 1222
Qy 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAAGACGAGAAGTTTACAGTACCTG 1737
Db 1223 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAAGACGAGAAGTTTACAGTACCTG 1282
Qy 1738 COTTATACAGGACACCTATACCTGATTAACCTTAAACGACAAATAA 1782
Db 1283 CGTTATACAGGACACCTATACCTGATTAACCTTAAACGACAAATAA 1327

RESULT 6

US-10-631-558-10
; Sequence 10, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002

; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
; US-10-631-558-10

Query Match 64.5%; Score 1150.2; DB 7; Length 1661;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1152, Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGTGTT 597
Db 183 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATGGTGTGTT 242
Qy 598 AGCGTTGCTGTTACTGTTGAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGGAAATC 657
Db 243 AGCGTTGCTGTTACTGTTGAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGGAAATC 302
Qy 658 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATATCA 717
Db 303 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGCTTTAAGTCCAAATATCA 362
Qy 718 AAACCAATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 777
Db 363 AAACCAATTTGCTACTGATAGTGGCGCGATGTACATAAACTTGAGAAAGCTGACTTACTA 422
Qy 778 AAGCTATTTCAAGAACAAATTTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTGC 837
Db 423 AAGCTATTTCAAGAACAAATTTGATCGCTAAACGTCACAGTAACGACGACTACTTTGAGTGC 482
Qy 838 ATTGATTTTCCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTCCTACTTTGCTGAC 897
Db 483 ATTGATTTTCCAGCGATGCAACCATTTACTGATCGAAACCGCAAGGTCCTACTTTGCTGAC 542
Qy 898 AAAGATGTTTCGGTAACTTTGCGGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 957
Db 543 AAAGATGTTTCGGTAACTTTGCGGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 602
Qy 958 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACAGCGAAATCTGTTGATGTG 1017
Db 603 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAAACAGCGAAATCTGTTGATGTG 662
Qy 1018 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 1077
Db 663 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTCCTCAA 722
Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCCGTTGACACCATCATCATCTCAAGAAATTA 1137
Db 723 GATACTAAGCTATTGAAAAACACTAGCTATCCGTTGACACCATCATCATCTCAAGAAATTA 782
Qy 1138 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTTGAC 1197
Db 783 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTTATGAAACGTTGAC 842
Qy 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 1257
Db 843 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 902
Qy 1258 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 903 TTTTACTTACCGTGTAAAAATCGGAAACAGCTTTATAGGATCAATAAAAAATCTGGTCTG 962

```
QY 1318 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 963 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1022
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACTGTTTCCACATCAATACGTT 1437
Db 1023 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACTGTTTCCACATCAATACGTT 1082
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db 1083 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1142
QY 1498 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTCTCTACAAATCTC 1557
Db 1143 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTCTCTACAAATCTC 1202
QY 1558 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATATACGATGAC 1617
Db 1203 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGATAGAGGATATACGATGAC 1262
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAGATGCTAGCTAT 1677
Db 1263 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1322
QY 1678 CATTTAGCCCTATGAT 1692
Db 1323 CATTTAGCCGCTGAT 1337
```

RESULT 7

```
US-10-631-558-9
; Sequence 9, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahn, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-9
```

```
Query Match 63.6%; Score 1134.2; DB 7; Length 1541;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 538 ATTGCTGGACCTGAGTGGCTGTAGACCGGTCCATCTGTCTCAACACGACCAATTTGGTTGTT 597
Db 33 ATAGCTGGTCCCTGAATGGCTACTAGATCGTCTCTCTGTAATAACAGCCAATTTGGTTGTT 92
```

```
QY 598 AGCGTTGCTGTACTGTTGAGGGGAGCGAATCAAGACATTAGTCTTTAAATTTTTTTGAAATC 657
Db 93 AGCGTTGCTGTACTGTTGAGGGGAGCGAATCAAGACATTAGTCTTTAAATTTTTTTGAAATC 152
QY 658 GATCTAACATCAGCAGCTGCTCATGGAGGAAGACAGAGCAGGCTTAAGTCCAAATCA 717
Db 153 GATCTAACATCAGCAGCTGCTCATGGAGGAAGACAGAGCAGGCTTAAGTCCAAATCA 212
QY 718 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAAGTGCATTACTA 777
Db 213 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAAGTGCATTACTA 272
QY 778 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCCACAGTAAACGACGACTACTTTGAGGTC 837
Db 273 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCCACAGTAAACGACGACTACTTTGAGGTC 332
QY 838 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAAACGCGAAGGCTACTTTGCTGAC 897
Db 333 ATTGATTTTGCAGCGATGCAACCAATTACTGATCGAAAACGCGAAGGCTACTTTGCTGAC 392
QY 898 AAAGATGTTCCGTTAACTTCCCGACCCAACTGTCCAAAGAAATTTTCTAAGCGGCAT 957
Db 393 AAAGATGTTCCGTTAACTTCCCGACCCAACTGTCCAAAGAAATTTTCTAAGCGGCAT 452
QY 958 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTGATGTG 1017
Db 453 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTGATGTG 512
QY 1018 GAATATATCTGTAAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCGGCTCTCAA 1077
Db 513 GAATATATCTGTAAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCGGCTCTCAA 572
QY 1078 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
Db 573 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 632
QY 1138 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC 1197
Db 633 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGAACGTCAC 692
QY 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 1257
Db 693 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 752
QY 1258 TTTACTTACCGTGTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 753 TTTACTTACCGTGTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 812
QY 1318 AATGAAGAAATAAACAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 813 AATGAAGAAATAAACAAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 872
QY 1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTTCCACATCAATACGTT 1437
Db 873 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAAACCTGTTTCCACATCAATACGTT 932
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 1497
Db 933 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 992
QY 1498 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTCTCTACAAATCTC 1557
Db 993 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGGCTAAACTCTCTACAAATCTC 1052
QY 1558 GATGCTTTTGGTATTATGAGTACTATACCTTAACTGGAAGATAGAGGATATACGATGAC 1617
Db 1053 GATGCTTTTGGTATTATGAGTACTATACCTTAACTGGAAGATAGAGGATATACGATGAC 1112
QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1677
Db 1113 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGATGCTAGCTAT 1172
QY 1678 CATTTAGCCCTATGAT 1692
```


Qy	388	AGAA	TTGGAGACACCTGGAGCAAGAGGATAATCAGAGAAACCTGCTCCAGTGCATCTGC	447
Db	1093	AGAA	TTGGAGACACCTGGAGCAAGAGGATAATCAGAGAAACCTGCTCCAGTGCATCTGC	1152
Qy	448	ACGGCAACGGCCGAGAGAGTGGAATGTGAGAGGCACACCTCTGTGCAGACCACATCG	507	
Db	1153	ACGGCAACGGCCGAGAGAGTGGAATGTGAGAGGCACACCTCTGTGCAGACCACATCG	1212	
Qy	508	AGCGGATCTGGCCCCCTTACCGATGTTTCGTAATTCCTGGACCTGAGTGGCTGTAGACCGT	567	
Db	1213	AGCGGATCTGGCCCCCTTACCGATGTTTCGTCGACCTGTTTACCAACCGAGCCTCACCCC	1272	
Qy	568	CCATCTGTCAACAACAGCCCAATTGGT	593	
Db	1273	CAGCCTCTCCCTATGGCCACTGTGT	1298	

```

RESULT 11
US-10-995-561-117
; Sequence 117, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117
; LENGTH: 7823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-117

```

RESULT 12
US-10-995-561-111
; Sequence 111, Application US/10995561

```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 7848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-111

```

```

RESULT 13
US-10-995-561-113
; Sequence 113, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 7935
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-113

```

Query Match	13.1%	Score 232.8;	DB 7;	Length 7935;
Best Local Similarity	79.8%	Pred. No. 4.5e-59;		
Matches 308;	Conservative	0;	Mismatches 32;	Indels 46;
				Gaps 1;

254	Qy	CCCATAGCTGAGAGTGTGTTTTCATCATCTGCTGGACCTTCTATATGGTTCGAGAAACG	313
913	Db	CCCATAGCTGAGAGTGTGTTTTCATCATCTGCTGGACCTTCTATATGGTTCGAGAAACG	972
314	Qy	-----TGGGAGAGGCGACG	327
973	Db	TGGGAGAGCGCCCTACCAAGGCTGGATGATGTGATCTTGCCTGGGAGAAAGCGACG	1032
328	Qy	GGACGCATCACTTGCACCTTCTAGAAATAGATGCCAACGATCAGGACACAAAGGACATCCTAT	387
1033	Db	GGACGCATCACTTGCACCTTCTAGAAATAGATGCCAACGATCAGGACACAAAGGACATCCTAT	1092
388	Qy	AGAAATTGGAGACACTGTGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC	447
1093	Db	AGAAATTGGAGACACTGTGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGC	1152
448	Qy	ACAGGCAACGCGCAGAGAGAGTGGAGTGTGAGAGGCACACTCTGTGCAGACCAATCG	507
1153	Db	ACAGGCAACGCGCAGAGAGAGTGGAGTGTGAGAGGCACACTCTGTGCAGACCAATCG	1212
508	Qy	AGCGGATCTGGCCCCCTTACCAGATGTTTCGTATTGCTGGACGTGAGTGGCTGTAGACCGT	567
1213	Db	AGCGGATCTGGCCCCCTTACCAGATGTTTCGTATTGCTGGACGTGAGTGGCTGTAGACCGT	1272
568	Qy	CCATCTGTCAACAAAGCCCAATTGGT	593
1273	Db	CAGCTCTCTCCCTATGGCCACTGTGT	1298

```

RESULT 14
US-10-995-561-108
; Sequence 108, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE,
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 7959
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-108

```

```

Qy 508 AGCGGATCTGGCCCCCTTACCAGATGTTCTGATTGCTGGACCTGAGTGGCTGCTAGACCGT 567
Db 1213 AGCGGATCTGGCCCCCTTACCAGATGTTCTGATTGCTGGACCTGAGTGGCTGCTAGACCGT 1272
Qy 568 CCATCTGTCACACACAGCCCAATTGGT 593
Db 1273 CAGCCTCTCCCTATGGCCACTGTGT 1298

RESULT 15
US-10-995-561-104
; Sequence 104, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 8013
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-104

Query Match 13.1%; Score 232.8; DB 7; Length 8013;
Best Local Similarity 79.8%; Pred. No. 4 5e-59;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

Qy 254 CCCATAGCTGAGAAGTGTGTTTGTATCATGCTGCTGGGACTTCCTATGTGTGTCGAGAAACG 313
Db 913 CCCATAGCTGAGAAGTGTGTTTGTATCATGCTGCTGGGACTTCCTATGTGTGTCGAGAAACG 972
Qy 314 -----TGGGAGAAGGCAGC 327
Db 973 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTATTCCTGGGAGAAGGCAGC 1032
Qy 328 GGAAGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 387
Db 1033 GGAAGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 1092
Qy 388 AGAATTGGAGACACCTGGAGCAAGAAGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 447
Db 1093 AGAATTGGAGACACCTGGAGCAAGAAGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
Qy 448 ACAGGCACCGCCGAGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGACAGACCAATCG 507
Db 1153 ACAGGCACCGCCGAGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGACAGACCAATCG 1212
Qy 508 AGCGGATCTGGCCCCCTTACCAGATGTTCTGATTGCTGGACCTGAGTGGCTGCTAGACCGT 567
Db 1213 AGCGGATCTGGCCCCCTTACCAGATGTTCTGATTGCTGGACCTGAGTGGCTGCTAGACCGT 1272
Qy 568 CCATCTGTCACACACAGCCCAATTGGT 593
Db 1273 CAGCCTCTCCCTATGGCCACTGTGT 1298

Search completed: February 1, 2006, 14:43:15
Job time : 331.469 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 7322.56 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-11
Perfect score: 1782
Sequence: 1 tgcgttcacgttcgtcgcg.....ataacccaacgacaaataa 1782

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	233.8	13.1	943	5	BX391752	BX391752 BX391752
2	232.8	13.1	451	1	AL603368	AL603368 DKFZp686C
C 3	232.8	13.1	465	1	AI095589	AI095589 qb24a08.x
4	232.8	13.1	621	7	CN482442	CN482442 hw20d08.y
5	232.8	13.1	626	7	CN419611	CN419611 170005313
6	232.8	13.1	765	3	BM715855	BM715855 UI-E-EU0-
7	232.8	13.1	7501	4	BC078656	BC078656 Homo sapi
8	232.8	13.1	7501	4	BC100030	BC100030 Homo sapi
9	232.8	13.1	7777	4	CR749316	CR749316 Homo sapi
10	232.8	13.1	7868	4	HSMB06653	HSMB06653 Homo sapi
11	232.8	13.1	7885	4	CR749281	CR749281 Homo sapi
12	232.8	13.1	8411	4	HSMB06992	HSMB06992 Homo sapi
C 13	231.2	13.0	548	6	CD613789	CD613789 Homo sapi
14	231.2	13.0	551	6	CD613788	CD613788 56022208H
15	231.2	13.0	861	1	AU140971	AU140971 AU140971
16	231.2	13.0	8121	4	CR749317	CR749317 Homo sapi
17	230.2	12.9	560	7	CN419585	CN419585 170005318
18	230.2	12.9	737	1	AU140993	AU140993 AU140993
19	229.2	12.9	427	7	CN419650	CN419650 170005315
20	229.2	12.9	452	7	CN419649	CN419649 170005315
21	229.2	12.9	547	2	BE009640	BE009640 PM4-BN017
22	229.2	12.9	580	3	BQ292415	BQ292415 PM0-AN008

23	229.2	12.9	899	5	BX327266	BX327266 BX327266
24	227.6	12.8	314	7	CN419631	CN419631 170005321
25	227.6	12.8	496	2	BG945197	BG945197 PM0-AN008
C 26	227.2	12.7	725	5	BQ574857	BQ574857 UI-H-E21-
27	227	12.7	734	1	AU140802	AU140802 AU140802
28	226.6	12.7	517	2	BF081716	BF081716 PM0-AN008
C 29	225.4	12.6	560	1	AI743013	AI743013 w85a12.x
30	223.8	12.6	393	1	AA376374	AA376374 EST8780
31	219.2	12.3	500	2	BF956982	BF956982 QV4-NN114
32	217.4	12.2	303	1	AA852090	AA852090 NHTBCae10
33	216.6	12.2	739	1	AL706215	AL706215 DKFZp686J
34	215	12.1	705	1	AU140789	AU140789 AU140789
35	214.6	12.0	901	5	BX327267	BX327267 BX327267
36	214	12.0	598	3	BQ340570	BQ340570 PM0-NN025
37	213.6	12.0	552	2	BG945246	BG945246 PM0-AN008
C 38	211.6	11.9	564	2	BF229278	BF229278 PM0-AN008
39	210.8	11.8	603	7	CN419658	CN419658 170005315
40	210.4	11.8	414	1	AA492032	AA492032 n955a12.8
41	210	11.8	400	3	BQ340575	BQ340575 PM0-NN025
42	209	11.7	577	3	BQ340604	BQ340604 PM0-NN025
C 43	206.8	11.6	367	3	BQ340606	BQ340606 PM0-NN025
44	203.6	11.4	454	8	W46530	W46530 zc32h07.r1
C 45	203	11.4	518	1	AI093548	AI093548 qb08h09.x

ALIGNMENTS

RESULT 1 BX391752/c
LOCUS BX391752 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CS0D1065Y122 3-PRIME, mRNA sequence.
ACCESSION BX391752 GI:30607701
VERSION BX391752.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1413.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/s=CS08AI025ZG09_CS02375_lkc=1413.r

FEATURES
source Location/Qualifiers
1. .943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CS0D1065Y122"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 13.1%; Score 233.8; DB 5; Length 943;

Best Local Similarity 80.0%; Pred. No. 2.4e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 45; Gaps 1;

Qy 254 CCCATAGCTGAGAAGTGTGTTTGTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 313
Db 778 CCCATAGCTGAGAAGTGTGTTTGTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 719

Qy 314 TG-----GGAGAGGCGAGCG 328
Db 718 TGGGAGAAGCCCTACCAAGCTGGATGGTAGATTGTACTTGGCTGGGAGAAGGCGAGCG 659

Qy 329 GAGCGATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTATA 388
Db 658 GAGCGATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTATA 599

Qy 389 GAAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGCA 448
Db 598 GAAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGCA 539

Qy 449 CAGGCAACCGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGACACACATCGA 508
Db 538 CAGGCAACCGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGACACACATCGA 479

Qy 509 CGCGATCTGCGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGTC 568
Db 478 CGCGATCTGCGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGTC 419

Qy 569 CATCTGTGCAACACAGCCCAATTGGT 593
Db 418 AGCCTCTCCCTATGCGCCACTGTGT 394

RESULT 2
AL603368
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKP2); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKF2p686C067) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p686C067"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match 13.1%; Score 232.8; DB 1; Length 451;
Best Local Similarity 79.8%; Pred. No. 3.7e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

Qy 254 CCCATAGCTGAGAAGTGTGTTTGTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 313
Db 32 CCCATAGCTGAGAAGTGTGTTTGTGATCATGCTGCTGGGACTTCCTATGTTGGTTCGGAGAAACG 91

Qy 314 -----TGGGAGAAGGCGAGCG 327
Db 92 TGGGAGAAGCCCTACCAAGCTGGATGGTAGATTGTACTTGGCTGGGAGAAGGCGAGCG 151

Qy 328 GAGCGATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT 387
Db 152 GAGCGATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT 211

Qy 388 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 447
Db 212 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 271

Qy 448 ACAGGCAACCGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGACACACATCG 507
Db 272 ACAGGCAACCGCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCGACACACATCG 331

Qy 508 AGCGGATCTGCGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGT 567
Db 332 AGCGGATCTGCGCCCTTCCACGATGTTCTGATTTGCTGGACCTGAGTGGCTCTAGACCGT 391

Qy 568 CCATCTGTCAACACAGCCCAATTGGT 593
Db 392 CAGCCTCTCCCTATGCGCCACTGTGT 417

RESULT 3
AI095589/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1697174"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -

AI095589 465 bp mRNA linear EST 05-OCT-1998
Q52408.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
IMAGE:1697174 3' similar to gb:X02761_cds1_FIBRONECTIN PRECURSOR
(HUMAN); mRNA sequence.
AI095589
AI095589.1 GI:3434565
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 465)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1871 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 446.
Location/Qualifiers
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1697174"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -

Genome Res. 6 (9), 791-806 (1996)
889548
Contact: Soares, ME
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu

Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

```

/obj_dir=/caxon:5000
/clone="UI-E-EJO-ahj-h-11-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev stage="fetal and adult"

```

```

/c/tissue_type="lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev stage="fetal and adult"

```

```
/rna_node=dnabv (viral technologies) (ii phage resistant)/
/clonelib="lib-E-EJ0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UII-E-EJ0 is a subtracted cDNA library constructed
```

/note="Organ: eye; vector: p1713-Fac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed

UI-E-500 is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed

0:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac.

stranded cDNA was ligated to an *scor* 1 adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (gt)₁₈ tail. The

vector. The oligonucleotide used to prime the cDNA synthesis of the first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The sequence tags for this library are: fetal eyes, AGATTACAGA; lens, CGATTACAGA; eye anterior segment, AGATTACAGA; lens, CGATTACAGA; eye anterior segment,

located between the NOL1 site and the Q11.8 tail. The sequence tags for this library are: fetal eyes, AGATACAGA; lens, CGATTACGGA; eye anterior segment, AATCGCGAT; optic nerve, CCATTAGTG; retina, CCGGG; Retina Foveal and Macular CTGCGGG and Choroid AGCA. This

AAATCCAGAT; IGUS, CGAATACAG; TAGT, eye anterior segment, AATCCGATG; optic nerve, CCATTAATG; retina, CCGGG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

foveal and macular, GICC; RFB and Choriola, ACCIA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

visual system, supported by National Eye Institute (NEI).

Similarity 13.1%; Score 232.8; DB 3; Length 765;
79.8%; Pred. No. 4.3e-55;

1 Similarity 79.8%; Pred. NO. 4.3e-55;
108; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

54 CCATAGCTGAGAAGTGTTTTGATCATGCTGCTGGGACTTCCATGTGTCGGAGAAACG 313

332 CCCATAGCTGAGAAGTGTTTTGATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACG 241

14 -----TGGGAGAAGGCAGC 327
|||

12 TGGGAGAAGCCCTACCAAGGCTGGATGATGTAGATTGTACTTGCCTGGGAGAAGGCAGC 301

388 GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 387

02 GGACGCATCACTTGCCTCTAGAAATAGTGAACGATCAGGACACAAGGACATCCTAT 361

388 AGAATTGGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 447

52 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 421

8 ACAGGCAACGGCCGAGGAGTGAAGTGTGAGGGCACCTCTGTGCAGACCACATCG 507
 2 ACAGGCAACGGCCGAGGAGTGAAGTGTGAGGGCACCTCTGTGCAGACCACATCG 481

222 ACAGGCAACGGCCGAGGAGTGGAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG 481

D8 AGCGGATCTGGCCCCCTTCACCGATGTTCTGATTGCTGGACCTGAGTGGCTGCTAGACCGT 567

[illegible]

Db 482 AGCGGATCTGCGCCCTTACCGAGTTCGTGACAGCTGTTTACCAACCGCAGCCTCACCCC 541

Qy 568 CCATCTGTCAACAACAGCCCAATTGGT 593

Db 542 CAGCCTCTCCCTATGCGCCACTGTGT 567

RESULT 7

BC078656

LOCUS

DEFINITION Homo sapiens cdna clone IMAGE:30347017, containing frame-shift errors

ACCESSION BC078656

VERSION BC078656.1 GI:50925326

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 7501)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Tohiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.W., Villalon,D.K., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 7501)

NIH MGC Project

Direct Submission

Submitted (02-AUG-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-x@mail.nih.gov

Tissue Procurement: Dr. Stefan Hansson

cDNA Library Preparation: Michael Brownstein / Ted Usdin

Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-ehgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAK Plate: 168 Row: 0 Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 47132556

This clone has the following problem: frame shifted.

Location/Qualifiers

1. .7501

/organism="Homo sapiens"

FEATURES

source

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30347017"

/tissue_type="Placenta, normal"

/clone_lib="NIH MGC_147"

/lab_host="DH10B"

/note="Vector: pBluescriptR"

ORIGIN

Query Match 13.1%; Score 232.8; DB 4; Length 7501;

Best Local Similarity 79.8%; Pred. No. 8.6e-55;

Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;

Qy 254 CCCATAGCTGAGAAAGTGTTCATCATCTGCTGGGACTTCTTATGTGGTGGGAAACG 313

Db 806 CCCATAGCTGAGAAAGTGTTCATCATCTGCTGGGACTTCTTATGTGGTGGGAAACG 865

Qy 314 -----TGGGAGAGGCGAGC 327

Db TGGGAGAGGCGCTACCAAGGCTGGATGATGGTAGATTGTACTTGGCTGGGAGAGGCGAGC 925

Qy 328 GGAGCGCATCTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 387

Db 926 GGAGCGCATCTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 985

Qy 388 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 447

Db 986 AGAATTGGAGACACCTGGAGCAAGAGGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 1045

Qy 448 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 507

Db 1046 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 1105

Qy 508 AGCGGATCTGCGCCCTTCCACGATGTTGTTATGCTGACCTGAGTGGCTGCTGACCGT 567

Db 1106 AGCGGATCTGCGCCCTTCCACGATGTTGTTATGCTGACCTGTTTACCAACCGCAGCCTCACCCC 1165

Qy 568 CCATCTGTCAACAACAGCCCAATTGGT 593

Db 1166 CAGCCTCTCCCTATGCGCCACTGTGT 1191

RESULT 8

BC100030

LOCUS

DEFINITION Homo sapiens cdna clone IMAGE:30343682, containing frame-shift errors

ACCESSION BC100030

VERSION BC100030.1 GI:71679847

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 7501)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Tohiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,

Schnarch, A., Schein, J.E., Jones, S.J. and Warra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 7501)
NIH MGC Project
Direct Submission
Submitted (01-AUG-2005) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 198 Row: j Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 47132556
This clone has the following problem: frame shifted.
FEATURES
source
1..7501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30343682"
/tissue_type="Placenta, normal"
/clone_lib="NIH_MGC_147"
/lab_host="DH10B"
/note="vector: pluescriptR"
ORIGIN
Query Match 13.1%; Score 232.8; DB 4; Length 7501;
Best Local Similarity 79.8%; Pred. No. 8.6e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAAGTGTCTTTTGATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACG 313
Db 806 CCCATAGCTGAGAAGTGTCTTTTGATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACG 865
QY 314 -----TGGAGAGGGCAGC 327
Db 866 TGGAGAGAGCCCTACCAAGGCTGGATGATGATGTAGATTGTACTTTCCTCGGGAGAGGCAGC 925
QY 328 GGACGCATCACTTGCATCTCTAGAAATAGATGCAACGATCAGGACACAAGACATCCTAT 387
Db 926 GGACGCATCACTTGCATCTCTAGAAATAGATGCAACGATCAGGACACAAGACATCCTAT 985
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 447
Db 986 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 1045
QY 448 ACAGGCCACCGCGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGCGAGACCAATCG 507
Db 1046 ACAGGCCACCGCGAGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGCGAGACCAATCG 1105
QY 508 AGCGGATCTGC3CCCTTACCAGATGTTTCGATTGCTCGACCTGAGTGGCTGTAGACCGT 567
Db 1106 AGCGGATCTGC3CCCTTACCAGATGTTTCGATTGCTCGACCTGTTTACCAACCGCAGCCTCACCCC 1165
QY 568 CCATCTGTCAACACAGCCAAATGGT 593

Db 1166 CAGCCTCTCCCTGATGCGCACTGTGT 1191

RESULT 9
 CR749316
 LOCUS
 DEFINITION Homo sapiens mRNA; cDNA DKF2p686K08164 (from clone DKF2p686K08164).
 ACCESSION CR749316
 VERSION CR749316.1 GI:51476361
 KEYWORDS HTC.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 7777)
 AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
 CONSRTM The German CDNA Consortium
 TITLE Direct Submission
 JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKF2p686K08164) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKF2p686K08164
 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
 FEATURES
 source
 1..7777
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKF2p686K08164"
 /tissue_type="uterus endothel, primary cell culture"
 /clone_lib="686 (synonym: hlcc3). Vector pSporti_Sfi; host DH10B; sites SfiIA + SfiIB"
 /dev_stage="adult"
 /note="fibronectin 1, differentially spliced"
 1..7777
 /gene="DKF2p686K08164"
 266..7069
 /gene="DKF2p686K08164"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAH18171.1"
 /db_xref="GI:51476362"
 /translation="MLRGGPGLLLAVLCGLTAVPSTGASKSKRRQAQKQVQSPQV
 VQSQKPCDNGKHQYIQINQWERTYLGNALVCTCYGSGRNCESPFAEETCFDKYT
 GNTYRGVDYTRPKDSMIWDCITCIGAGRGISCTIANRGHGQSQYKIGTWRRPHT
 GGYMLRCVCLGNGKQWTKPIAEKCFDHAAGTSYVVGTEKPYQGMWVDCITCLGE
 QSRTICTSRNCRNDQDTSTSYRIGDTWSKDNRGNLQICITGNGRGWKCEHTNSV
 GTSITGSGCTDRAAYVQPPHPPYPCGVCHCVTSVSGVSGWMLKTQGNQKMLCT
 CLNGVSCQETAATVQYNGNSGKPCVLPTFYNGRTTFYCTTEGRDQHLWCSTTNSI
 EQDQTSFCFTHVLTVYGRGNSGKALCHFFPLNNYNTDCTSEGRDNMKWCGTTC
 NYADQKFGFCGMAAEHETCTTNGVMYRIGDQWQHDHGMHMRCTCVGNGVSGSTEL
 IYASQLRDOCIVDDITVYNDTPHKEHGHMLNCTFCQGRGRWKCDDPVDQCDQSET
 GTFPYIGDSWEKXVHGVRQCYCYGRGIGWHCQPLQTYTPSSSGPVEVFTETPSQPN
 SHPIQWNAQPSHSKTYILWRPNSVGRWKEATIPGLHNSYTIKGLKPGVVEVTEGQI
 STQQYGHQVETFPDFTTSTSTPTVNTVGTTFPSPLVATSVESVTEITASSVFWVS
 VASADTSQTRFVEYLSDESDQYLDLPSTASVNI PDLPGKRTIVNYQISJEDGE
 QSLILSTGOTFADPDPDQVDDTSYIVRMSQPAQITGYRI VVSPVSGSSTEL
 NLPETANSYTLSDQPGVQYNIITVAYEENQESTPVYIQETTGTPSDTPVSPRLQ
 FVEYTDVKYTIWMTPEASATGYRVDVIPNLPGHGQRPLISRNTFAEYVGLSPGV
 YYPKFAVSHGRESKPLQAQITKLDAPNLQFVNEDSTVILVRVTPPAETITGYRLT
 CLNGVSCQETAATVQYNGNSGKPCVLPTFYNGRTTFYCTTEGRDQHLWCSTTNSI
 EQDQTSFCFTHVLTVYGRGNSGKALCHFFPLNNYNTDCTSEGRDNMKWCGTTC
 NYADQKFGFCGMAAEHETCTTNGVMYRIGDQWQHDHGMHMRCTCVGNGVSGSTEL
 IYASQLRDOCIVDDITVYNDTPHKEHGHMLNCTFCQGRGRWKCDDPVDQCDQSET
 GTFPYIGDSWEKXVHGVRQCYCYGRGIGWHCQPLQTYTPSSSGPVEVFTETPSQPN
 SHPIQWNAQPSHSKTYILWRPNSVGRWKEATIPGLHNSYTIKGLKPGVVEVTEGQI
 STQQYGHQVETFPDFTTSTSTPTVNTVGTTFPSPLVATSVESVTEITASSVFWVS
 VASADTSQTRFVEYLSDESDQYLDLPSTASVNI PDLPGKRTIVNYQISJEDGE
 QSLILSTGOTFADPDPDQVDDTSYIVRMSQPAQITGYRI VVSPVSGSSTEL
 NLPETANSYTLSDQPGVQYNIITVAYEENQESTPVYIQETTGTPSDTPVSPRLQ
 FVEYTDVKYTIWMTPEASATGYRVDVIPNLPGHGQRPLISRNTFAEYVGLSPGV
 YYPKFAVSHGRESKPLQAQITKLDAPNLQFVNEDSTVILVRVTPPAETITGYRLT
 CLNGVSCQETAATVQYNGNSGKPCVLPTFYNGRTTFYCTTEGRDQHLWCSTTNSI
 EQDQTSFCFTHVLTVYGRGNSGKALCHFFPLNNYNTDCTSEGRDNMKWCGTTC
 NYADQKFGFCGMAAEHETCTTNGVMYRIGDQWQHDHGMHMRCTCVGNGVSGSTEL
 IYASQLRDOCIVDDITVYNDTPHKEHGHMLNCTFCQGRGRWKCDDPVDQCDQSET
 GTFPYIGDSWEKXVHGVRQCYCYGRGIGWHCQPLQTYTPSSSGPVEVFTETPSQPN
 SHPIQWNAQPSHSKTYILWRPNSVGRWKEATIPGLHNSYTIKGLKPGVVEVTEGQI
 STQQYGHQVETFPDFTTSTSTPTVNTVGTTFPSPLVATSVESVTEITASSVFWVS
 VASADTSQTRFVEYLSDESDQYLDLPSTASVNI PDLPGKRTIVNYQISJEDGE
 QSLILSTGOTFADPDPDQVDDTSYIVRMSQPAQITGYRI VVSPVSGSSTEL
 NLPETANSYTLSDQPGVQYNIITVAYEENQESTPVYIQETTGTPSDTPVSPRLQ
 FVEYTDVKYTIWMTPEASATGYRVDVIPNLPGHGQRPLISRNTFAEYVGLSPGV
 YYPKFAVSHGRESKPLQAQITKLDAPNLQFVNEDSTVILVRVTPPAETITGYRLT
 CLNGVSCQETAATVQYNGNSGKPCVLPTFYNGRTTFYCTTEGRDQHLWCSTTNSI
 EQDQTSFCFTHVLTVYGRGNSGKALCHFFPLNNYNTDCTSEGRDNMKWCGTTC
 NYADQKFGFCGMAAEHETCTTNGVMYRIGDQWQHDHGMHMRCTCVGNGVSGSTEL
 IYASQLRDOCIVDDITVYNDTPHKEHGHMLNCTFCQGRGRWKCDDPVDQCDQSET
 GTFPYIGDSWEKXVHGVRQCYCYGRGIGWHCQPLQTYTPSSSGPVEVFTETPSQPN
 SHPIQWNAQPSHSKTYILWRPNSVGRWKEATIPGLHNSYTIKGLKPGVVEVTEGQI
 STQQYGHQVETFPDFTTSTSTPTVNTVGTTFPSPLVATSVESVTEITASSVFWVS
 VASADTSQTRFVEYLSDESDQYLDLPSTASVNI PDLPGKRTIVNYQISJEDGE
 QSLILSTGOTFADPDPDQVDDTSYIVRMSQPAQITGYRI VVSPVSGSSTEL
 NLPETANSYTLSDQPGVQYNIITVAYEENQESTPVYIQETTGTPSDTPVSPRLQ
 FVEYTDVKYTIWMTPEASATGYRVDVIPNLPGHGQRPLISRNTFAEYVGLSPGV
 YYPKFAVSHGRESKPLQAQITKLDAPNLQFVNEDSTVILVRVTPPAETITGYRLT
 CLNGVSCQETAATVQYNGNSGKPCVLPTFYNGRTTFYCTTEGRDQHLWCSTTNSI
 EQDQTSFCFTHVLTVYGRGNSGKALCHFFPLNNYNTDCTSEGRDNMKWCGTTC
 NYADQKFGFCGMAAEHETCTTNGVMYRIGDQWQHDHGMHMRCTCVGNGVSGSTEL
 IYASQLRDOCIVDDITVYNDTPHKEHGHMLNCTFCQGRGRWKCDDPVDQCDQSET
 GTFPYIGDSWEKXVHGVRQCYCYGRGIGWHCQPLQTYTPSSSGPVEVFTETPSQPN
 SHPIQWNAQPSHSKTYILWRPNSVGRWKEATIPGLHNSYTIKGLKPGVVEVTEGQI
 STQQYGHQVETFPDFTTSTSTPTVNTVGTTFPSPLVATSVESVTEITASSVFWVS
 VASADTSQTRFVEYLSDESDQYLDLPSTASVNI PDLPGKRTIVNYQISJEDGE
 QSLILSTGOTFADPDPDQVDDTSYIVRMSQPAQITGYRI VVSPVSGSSTEL
 NLPETANSYTLSDQPGVQYNIITVAYEENQESTPVYIQETTGTPSDTPVSPRLQ
 FVEYTDVKYTIWMTPEASATGYRVDVIPNLPGHGQRPLISRNTFAEYVGLSPGV
 YYPKFAVSHGRESKPLQAQITKLDAPNLQFVNEDSTVILVRVTPPAETITGYRLT
 CLNGVSCQETAATVQYNGNSGKPCVLPTFYNGRTTFYCTTEGRDQHLWCSTTNSI
 EQDQTSFCFTHVLTVYGRGNSGKALCHFFPLNNYNTDCTSEGRDNMKWCGTTC
 NYADQKFGFCGMAAEHETCTTNGVMYRIGDQWQHDHGMHMRCTCVGNGVSGSTEL
 IYASQLRDOCIVDDITVYNDTPHKEHGHMLNCTFCQGRGRWKCDDPVDQCDQSET
 GTFPYIGDSWEKXVHGVRQCYCYGRGIGWHCQPLQTYTPSSSGPVEVFTETPSQPN
 SHPIQWNAQPSHSKTYILWRPNSVGRWKEATIPGLHNSYTIKGLKPGVVEVTEGQI
 STQQYGHQVETFPDFTTSTSTPTVNTVGTTFPSPLVATSVESVTEITASSVFWVS
 VASADTSQTRFVEYLSDESDQYLDLPSTASVNI PDLPGKRTIVNYQISJEDGE
 QSLILSTGOTFADPDPDQVDDTSYIVRMSQPAQITGYRI VVSPVSGSSTEL
 NLPETANSYTLSDQPGVQYNIITVAYEENQESTPVYIQETTGTPSDTPVSPRLQ
 FVEYTDVKYTIWMTPEASATGYRVDVIPNLPGHGQRPLISRNTFAEYVGLSPGV
 YYPKFAVSHGRESKPLQAQITKLDAPNLQFVNEDSTVILVRVTPPAETITGYRLT
 CLNGVSCQETAATVQYNGNSGKPCVLPTFYNGRTTFYCTTEGRDQHLWCSTTNSI
 EQDQTSFCFTHVLTVYGRGNSGKALCHFFPLNNYNTDCTSEGRDNMKWCGTTC
 NYADQKFGFCGMAAEHETCTTNGVMYRIGDQWQHDHGMHMRCTCVGNGVSGSTEL
 IYASQLRDOCIVDDITVYNDTPHKEHGHMLNCTFCQGRGRWKCDDPVDQCDQSET
 GTFPYIGDSWEKXVHGVRQCYCYGRGIGWHCQPLQTYTPSSSGPVEVFTETPSQPN
 SHPIQWNAQPSHSKTYILWRPNSVGRWKEATIPGLHNSYTIKGLKPGVVEVTEGQI
 STQQYGHQVETFPDFTTSTSTPTVNTVGTTFPSPLVATSVESVTEITASSVFWVS
 VASADTSQTRFVEYLSDESDQYLDLPSTASVNI PDLPGKRTIVNYQISJEDGE
 QSLILSTGOTFADPDPDQVDDTSYIVRMSQPAQITGYRI VVSPVSGSSTEL
 NLPETANSYTLSDQPGVQYNIITVAYEENQESTPVYIQET

VEVVYTVLQVLDGQERDAPVUNKVYVTPSPNTNHLRANPDVGVTVSWERSSTTPDII
 GYRVTTPVNGQOQNSLEBVHQAQSSCTSDNLSGPLEYVNVVTVVKKDKESVPISDI
 IIPVQLTDSJDSJGLRWMTSPNTTQVAVPNTLFRFTNGDQVTVVWVWVWVWVWV
 GIYTVLGLPGLDVIDSVLLINGESAPNTTQVAVPNTLFRFTNGDQVTVVWVWVWVWV
 APPDSIDLNTNGLFVSPKYNEDVBAELSIPSDNAVNTNLLPSTVYVSVSVSYEQH
 ESTLRGRQKTLGSDPTGSDFTDSTANSFHWIAPRATITGYRIHHHPHSGRPRE
 DRVPHSNRITLNLNTPCTEYVSVIYVANSRGENSPLLIQOQSVTVVPRDLVAAVPA
 TSLLIWDAPNAVYRYITVYETGCGNSPQSVFTVPKGSATISLPGKPCVDITVYV
 AVTRGDSPASAKPIISNRYETIDKPSQVOTVDQDINSIVKVLPSGSPGVTVYRV
 PKNGEPKTKTACGDQNTMEIGLOPVTYVYVVAQNPSESGOPLVOTVATVIPAP
 TDLKTOVTPKTSNAQTPVNVOLTGTVYVVPVPEKTPMKENLAPDSSVVVVGSLM
 VAKVEYSVALKOTLTSRPAQVTVTVLNVSPRPRVATDATETITISWRTKETI
 TGQVDVAPANGQPIQKRIIPDVRKSYTITLQGPQGYDKIYLLTNLNDKSSPVVIDA
 STADIVANPFLATPNTIIVKQVNSKPEIGRKTGQALQOTTISWAPQDTEYIIS
 ATITGLBEPETTYIVTALKNQKSEPIGRKTGQALQOTTISWAPQDTEYIIS
 CHVPDTEEPLOPVYVTSXATITGLTGRATNTIIVALKQOQRHAREVTVYVTS
 VNEGLNQPDTKSPDPVTVSHYVAGVDEWMSSESGKPLCQLGPHGFCRCDSSRWK
 HDNGVNTIGISKBDPQVGSNGQMSCTCLNGKHGEPKCDHATCYDQGTGYHVGEQWV
 KGLYALICSCCTFGQGRQVNDCCRPRGEPTEBGTGQSVNQSVQRVHQRTNVNVC
 PIETCEMLDVAQDSDSPR"

ORIGIN

Query Match	13.1%;	Score 232.8;	DB 4;	Length 7777;
Best Local Similarity	79.8%;	Prod. No. 8.7e-55;		
Matches 308;	Conservative	0; Mismatches 32;	Indels 46;	Gaps 1;
254	CCCATAGCTGAGAAAGTGTTTGGATCATCTCTGGGACTTCCTATGTGTGGGAAACG	313		
b				
806	CCCATAGCTGAGAGTGTTTGGATCATCTCTGGGACTTCCTATGTGTGGGAAACG	865		
314	-----TGGGAGAAAGGCAGC	327		
Y				
866	TGGGAGAAAGCCCTACCAAGGCTGGATGATGTTAGATTGTACTTGCCTGGGAGAAAGCAGC	925		
b				
328	GGATCCGATCATTGACATCTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT	387		
Y				
926	GGACCGATCATTGACATCTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAT	985		
b				
388	AGAAATTGGAGACACCTCGGAGCAAGAAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	447		
Y				
986	AGAAATTGGAGACACCTCGGAGCAAGAAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	1045		
b				
448	ACAGGACACGGCCGAGGAGTGGAAGTGTGAGAGGCACACCTCTGTGTGAGACCAACATCG	507		
Y				
1046	ACAGGCAACGGCCGAGGAGTGGAAGTGTGAGAGGCACACCTCTGTGTGAGACCAACATCG	1105		
b				
508	ACCGGATCTGGGCCCTCTCACCGATCTTCGTAATTGCTGGACCTGAGTGCCTAGACCGT	567		
Y				
1106	ACCGGATCTGGGCCCTCTCACCGATCTTCGTAATTGCTGGACCTGAGTGCCTAGACCGT	1165		
b				
568	CCATCTGTCAAACACAGCCAAATTGTT	593		
Y				
1166	CAGCCTCTCCCTATGGCCACTGTGT	1191		

RESULT 10	HSM806653	7868 bp	linear	HTC 20-JAN-2005
LOCUS	HSM806653			
DEFINITION	Homo sapiens mRNA; cDNA DKF2p686M04163 (from clone DKF2p686M04163).			
ACCESSION	BF640608			
VERSION	BF640608.1			
KEYWORDS	HTC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1. (bases 1 to 7868)			
AUTHORS	Bloecher,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C., Oanger,A., Fobo,G., Han,M. and Wiemann,S.			
CONSTRM	The German cDNA Consortium			
TITLE	Direct Submission			
JOURNAL	Submitted (20-JAN-2005) MIPS, Ingolstaedter Landstr.1. D-85764			

COMMENT

FEATURES

gene

CDS

Neuhberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686M04163) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M04163>
Further information about the clone and the sequencing project is available at <http://mips.gst.de/projects/cdna/>

```

1. .7868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686M04163Q"
/db_xref="taxon:9606"
/clone="DKFZp686M04163"
/tissue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlccc3). Vector psport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin precursor"
1. .7868
/gene="DKFZp686M04163"
267.. .7157
/gene="DKFZp686M04163"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45714.1"
/db_xref="GI:34364617"
/db_xref="GOA:G6N0A6"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"
/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/TrEMBL:G6N0A6"

```

/translation="MLRGPGFLLLLALGLGTA
 VSGQKPGDNHGRYINQWVOTJGNALVCYTGSGRSGHGGSCPEAKETCFDKYT
 NTRVYGDTYPRKDSIMWDCITGAGRGRTSCTIANRCHGEGSQSYGIDTWRRPHET
 GYMLCECVLGNKGWETKPIAEKCPDHAAGTSYVGETWEKPYQGMWMDCTCLGE
 GSGRICTCSNRKQNDQTRTSYRIGDTPWKKDNRGNLQOCITGCHGKWCERKHTSV
 QTTSSGSGPFTDRAAVTQPHQPHQYGHCVCTGDSVYSGMGLKTCQGNKMLCT
 CLNGVSQCEHTAVTQTVGNSNGSFCVLPFTYNGRTFYSTCTEGRQDGHLCWCTSTNY
 EDOQYGFCTDVTQJVGNSGNCALCHFPFLYNNHYDTCTDSGRDNKMCWCTGT
 NYDAQKFCPCMAHEECITINEGVRIRIGDMDQFGHGHMMRCTCVNGRGSEWT
 IYASQLRQDCIVDAITVYNNDTFKRHEEGHMLKCTQGMGRKWCDOQDOQDSET
 GTFYQIGDSWEKHYGVRYQCKYCRGIGEHWCQPLQTYPSSSGPVEVPIETPESQPN
 SHYIQMNAQVPSHLSKYILLRWRPNKGRWKEATCPHLNLSYITIKGLPGVYVEGQL
 SIQOQHQEPTSHSTPTVNTSVTGETTPEFLPVAITSVETIATSSVFW
 VASADTQVSGFRVEYLEBEGDEPQDLTPSTAYSITPDLPKRYIYVNYQISDEGE
 QSLTSLQSTAPADPEPQVDQDTPYIVWRVSQRAQITGYRIVYSPVSGSSTEL
 NLPTANSVTLSDLOPGVOYNIITVAVEENOESTPVLQOETTGTPRSDTPVPSRDLQ
 FVEYTDVNTVIMTPPESATGYRVDVTPVNLPCBQGRQLPIVRNTPAEPVGLSGPVT
 YIFKVFVSHGRMSTQTKKDDAFNLQFVNETISLVRNTPPRAQITGYKLT
 VGLTRGQRYNGVPSVKYPLRNLQPSAYTISVLSVAKGNQSPKATGVTTLPQ
 STTPYNNETVETIYITPAPRIGFLKGYRQSPGABREPTSDSGSIVVSGTLPG
 VEYVITVQLRQDGERADYKVKVUTPLSPITLXANPDGTVLWSBERSTTPDIT
 GYRITTTTPQOQNGSLVEVHADQSSCTPNLSPGLEVSVYVATKDKESVPISDT
 IIPAVPPTDLRFNIGPDTRVWAPPSIDLTPLYRSPVKNEEDVAELSISPSD
 NAVUTMLLPTEYSVSVSVYGEHSTPLKGRKIQIDSGPTGDSIDFSLTNSFTVHM
 IAPRATITGYRIHHPSHGRREDRYPHGRNSITLNLTPGTYSVYIVALNCRGE
 SPILLIGQOSTVSDPRLEVVAAPTSLIISWDAPAVTRYRITGYTGSGNSPQVEF
 TPWGSATSTISGLKPGYITIVVATVGRGDSPASSPKISINRYATEIKPSQOMVTD
 VQNGSIKVMKLPSSSPVTVYRITTPKNGPGTKTKTAGPDQETMEGLQITVEVY
 SVYAQNSGESOPVLAQVATITIPATDLKFQTYTISLAQTPWPNQLTGYRVRTP
 KEKTPWKEINLAPDSSSVVSGMLVATIKEVSVTALMDITLRPAQVNTYLENVS
 PRARVATDTEITILISWRTKTEITGFQDVAANPLRTPIQRTIKLPSVRSYITGLQ
 PGTDYKILYTLNDNARSPPYI DASTADIANPLRFLATPNSLKLPSVQPPRARITG
 YIYKYEKSPGSPRPPRGVATEATIGLEPGTEYITVYALKNKQSPBILGRKK
 TDLIPOLVITPHNIPHGRILQVPSVQKTPVWTHGVDTGNGIOLISGQOOPSQV
 T

Db 966 GGAAGCATCTCTGCACTTCTAGAAAATAGATGCAACGATCAGGACACAAAGGACATCTCTAT 1025
Qy 388 AGAATTGGAGACACCTGGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCGC 447
Db 1026 AGAATTGGAGACACCTGGAGCAAGAGTAATCGAGAAACCTGCTCCAGTGCATCGC 1085
Qy 448 ACAGGCAACCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGCAGACCAATCG 507
Db 1086 ACAGGCAACCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGCAGACCAATCG 1145
Qy 508 AGCGGATCTGCGCCCTTCCAGGATGTTCTGATTCCTGCACTGAGTGGCTCTAGACGT 567
Db 1146 AGCGGATCTGCGCCCTTCCAGGATGTTCTGATTCCTGCACTGAGTGGCTCTAGACGT 1205
Qy 568 CCATCTGTCAACACAGCCCAATTGGT 593
Db 1206 CAGCTCTCTCCCTATGGCCACTGTGT 1231

RESULT 12
HSM806992
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
HSM806992 8411 bp mRNA linear HTC 23-FEB-2005
Homo sapiens mRNA; cDNA DKFZp68601166 (from clone DKFZp68601166).
BX640875
BX640875.1 GI:34365169
HTC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 8411)
Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,
Osanger, A., Fobo, G., Han, M. and Wiemann, S.
The German CDNA Consortium
Direct Submission
Submitted (22-FEB-2005) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GfB (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp68601166) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp68601166
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
Location/Qualifiers
1..8411
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp68601166Q"
/db_xref="taxon:9606"
/clone="DKFZp68601166"
/tissue_type="endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin 1"
1..8411
/gene="DKFZp68601166"
268..7701
/gene="DKFZp68601166"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA545932.1"
/db_xref="GI:34365170"
/db_xref="GOA:Q6MZU5"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"

gene
CDS

/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/TreMBL:Q6MZU5"
/translation="MLRPGPGGLLLAVLCIGTAVPSTGASKSKRQAQMVQPSVPA
VSQKPGCTDNGKHQIINQWERTYLLNALVCTCYGSGRGNFNCESKPRABETCFDKYT
GNTVRVGTIERPKDSMTDCTCIGAGRICTIANRCHGGQSYKIGDWRRRPHET
GYMLCEVCLNGKEWTKCPAIEKCFDHAAGTSYVVGTEWKEPKQGMWDTCTCLGE
GSRITCTSRNRCNDQDRTSYIGDTSKKNRNLQICITGNRGEWKECHERTSV
QTSSTSGSPFTDRAAVYQPPHPPYCHCVTDSGVVYVGMQMLTQGNKMLCT
CLNGVSCQETAVTQYCGNSNGEPCVLPFTYNGRTFYFSCCTTEGQDGLHWCSTTSNY
EQDKYSCTDHTVLVQTRGNSNGLCHFPFLYNNHNTDCTSGRDNMKNMCCGTTQ
NYDADQKFGFCMAAHEBICITNEGMYRIIGDQDKQDMHMRCTCVGNGSGWTC
IAYSLQRDQICVDITVYNDTFHRRHEGMLNCTCFQGRKRWKCDPVDQCSBT
GTFYIGDSWEKYVHGVYQVCYGRGIGEWHCOPLOTYPSSSGPVEVPIETESQPN
SHPIQWNAPOPSHLSKYLRLWRPKNSVGRWKEATIPGLNSVTIKGLKPGVYEQILI
SIQYGHQEVTRFPTTSTPTVTSNTVGETTTPFSELVATESVTIASSFVSW
VSADTVSGFRVEYELSEBDEPQYLDLPSTATSVNIPLLPGRYIYVNVQISDEG
QSLILSTSQTTPADPPTVDVDDTSIVVWSRPOAPITGYRIYVSPSGESTEL
NLPTANSVTLSDLPQGVQYNIITIVAVENQESTPVIIQETGTPRSDTVPSPRDLQ
FVEVTVKVTIMWTPPESAVTGYRVDVI PVNLPGHGORLPI SRNTFAEVTGLSPGT
YFKVFAVSHGSKESKPLTAQOTTKLDAPTNLQFVNETSIVLVRWTPPRAQITGLRLT
VGLTRGQPRQINVPSPVSKYPLRLNPASEYTVSLVAIKGNQSPKATGVFTTLQPG
SSIPPYNTEVTETITVITPAPRIGFKLVGPSQSGAPREVTSDSGSIVVSGLTGP
VEYVYITQLRDGQERDAPIVKVVPLSPPTNLHLNAPDTGLVTSWERSITPDIT
GYRITPTTNGOQNSLREVQADQSSCTFDNLSPGLEVNVVTVKDDKESVPLSDT
IIEVPOLTDLSFVDITDSSI GLRWTPLNSSTIIGYRITVVAAGEGIIPIFEDFVDSV
GYVTGLPGIDVDISVITLINGSAPTTLTQQTAVPPPTDLKFTNIGDPTMKVTW
APPSIDLTNLFVSPVKNEDVABLSISPSDNAVLNTLLPGTYEYVSVSVSEYEH
ESTPLRGRTGLSDPTGDFDITANSFTVHMTAPRATITGYRIHHPHFSGRPRE
DRVPHSRNLTNLTPGTYEYVSVI VALNGREESPLLIQOOSTVSDVDRDLVVAATP
TSLIISWDAPATVYRYITGYETGNSPVQOBFVPGSKSTATISGLKPGVDYITTVY
AVTGRGSPASSKPI SINRYEIDKPSQOMVTDQDINSISVKNLPSSPVTVGYRTT
PXNGFTTKITAGPDQETMIEGLQPTVEYVSVIAQNPSSGSQLVQVATVINDRP
KGLATDVDDSIKIAMESQGVSRVRYTVSSPEDGHELFAPDGEDFAELRGLR
PGSEYTVSVVALHDMESQPLIGTQSTAPATDLKFTQVTSLSAQMTVPNQVLTG
YRVVTPKEKTGPKMEINLAPDSSVSVGLTVATKYEVSVALKDITSTSPAQCVLT
TLENVSPRRARVTDATETTTISWRTKETITCGVDVAPANGOTPIQRTIKPDRS
TYLNGLPQTDYKIYLLNDNARSPPVIDASTAIDAPSNLRLATTPNLSLLWQSP
PRATIGYIIKIKPSPREVPRPRPTEATITGLEPGTYIYVIALKNNKQSE
PLIKRKTDELQVLTPHNLHGEILDVSTVQKTFVTHPGVDTGNGQLGDTGSG
QVPSGQQMIPEEHGFRRTPTPTPIRHRPRYPVNVGIEI OI GHPREDVDVHLY
PHGLPNASTGOEALSOITISWAPFODTSEYIISCHPVGTDEEPLOFRVPGTSTA
TLGLTRGATYNIIVEALKOORHKVREVVTVGNSVNEGLNQYTDSCFPDPTVSHY
AVGEWERSSESGFKLJCQLGSGHGFRCSSRWCHDNGVNIKGEKWDGNGQM
MSTCLNGNGEKFCDPHEATCYDDGTYHVGEQWQKYLEGAI CSCTCFGQGRMCRD
NCRPRGGEPSPEGTGGSYNOYSORYHQTNTNTVNYPIECFMPDLVQADREDSRE"

ORIGIN
Query Match 13.1%; Score 232.8; DB 4; Length 8411;
Best Local Similarity 79.8%; Pred. No. 8.9e-55;
Matches 308; Conservative 0; Mismatches 32; Indels 46; Gaps 1;
Qy 254 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGAGCTTCTCTATGTGCTGGAGAAACG 313
Db 808 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGAGCTTCTCTATGTGCTGGAGAAACG 867
Qy 314 -----TGGGAGAAAGGCGACG 327
Db 868 TGGGAGAAAGGCGCTACCAAGCTGGATGATGCTAGATTGTA CTCTGCTGGAGAGGCGACG 927
Qy 328 GGACGCACTCATTGCATCTTAGAAATAGATGCAACGATCAGGACACAGGACATCTCAT 387
Db 928 GGACGCACTCATTGCATCTTAGAAATAGATGCAACGATCAGGACACAGGACATCTCAT 987
Qy 388 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGGAAACCTGCTCAGTGCATCTGC 447
Db 988 AGAATTGGAGACACCTGGAGCAAGAGATATCGAGGAAACCTGCTCAGTGCATCTGC 1047
Qy 448 ACAGGCAACCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGACCATCG 507
Db 1048 ACAGGCAACCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGTCAGACCATCG 1107

QY 508 AGCGATCTGGCCCTTACCGATGTCGTATTTGCTGGACCTGAGTGGCTGTAGACCGT 567
Db 1108 AGCGATCTGGCCCTTACCGATGTCGTATTTGCTGGACCTGAGTGGCTGTAGACCGT 1167
QY 568 CCATCTGTCAACACAGCCCAATTGGT 593
Db 1168 CAGCCTCTCCCTATGCGCCACTGTGT 1193

RESULT 13
CD613789/c
LOCUS 568 bp mRNA linear EST 12-JAN-2004
DEFINITION 56022208J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD613789
VERSION CD613789.1 GI:40262053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 548)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
PUBMED 15203218
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1. .548
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 13.0%; Score 231.2; DB 6; Length 548;
Best Local Similarity 79.5%; Pred. No. 1.1e-54;
Matches 307; Conservative 0; Mismatches 33; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAAGTGTGATCATGCTGCTGGGACTTCTATGTGTCGAGAAACG 313
Db 440 CCCATAGCTGAGAAGTGTGATCATGCTGCTAGGACTTCTATGTGTCGAGAAACG 381
QY 314 -----TGGGAGAGGCGC 327
Db 380 TGGGAGAGCCCTTACCAAGCCGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCGC 321
QY 328 GGACGCATCACTTGCATCTTAGAAATAGATGCAACGATCAGGACACAGGACATCTAT 387
Db 320 GGACGCATCACTTGCATCTTAGAAATAGATGCAACGATCAGGACACAGGACATCTAT 261
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTGCATCTGC 447
Db 260 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTGCATCTGC 201
QY 448 ACAGGCAACGGCGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCACTATCG 507
Db 200 ACAGGCAACGGCGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCACTATCG 141
QY 508 AGCGATCTGGCCCTTACCGATGTCGTATTTGCTGGACCTGAGTGGCTGTAGACCGT 567
Db 140 AGCGATCTGGCCCTTACCGATGTCGTATTTGCTGGACCTGAGTGGCTGTAGACCGT 81
QY 568 CCATCTGTCAACACAGCCCAATTGGT 593
Db 80 CAGCCTCTCCCTATGCGCCACTGTGT 55

RESULT 14
CD613788
LOCUS 551 bp mRNA linear EST 12-JAN-2004
DEFINITION 56022208H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD613788
VERSION CD613788.1 GI:40262052
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
PUBMED 15203218
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1. .551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 13.0%; Score 231.2; DB 6; Length 551;
Best Local Similarity 79.5%; Pred. No. 1.1e-54;
Matches 307; Conservative 0; Mismatches 33; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAAGTGTGATCATGCTGCTGGGACTTCTATGTGTCGAGAAACG 313
Db 109 CCCATAGCTGAGAAGTGTGATCATGCTGCTAGGACTTCTATGTGTCGAGAAACG 168
QY 314 -----TGGGAGAGGCGC 327
Db 169 TGGGAGAGCCCTTACCAAGCCGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCGC 228
QY 328 GGACGCATCACTTGCATCTTAGAAATAGATGCAACGATCAGGACACAGGACATCTAT 387
Db 229 GGACGCATCACTTGCATCTTAGAAATAGATGCAACGATCAGGACACAGGACATCTAT 288
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTGCATCTGC 447
Db 289 AGAATTGGAGACACCTGGAGCAAGAGGATPAATCGAGGAAACCTGCTCCAGTGCATCTGC 348
QY 448 ACAGGCAACGGCGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCACTATCG 507
Db 349 ACAGGCAACGGCGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCACTATCG 408
QY 508 AGCGATCTGGCCCTTACCGATGTCGTATTTGCTGGACCTGAGTGGCTGTAGACCGT 567
Db 409 AGCGATCTGGCCCTTACCGATGTCGTATTTGCTGGACCTGAGTGGCTGTAGACCGT 468
QY 568 CCATCTGTCAACACAGCCCAATTGGT 593
Db 469 CAGCCTCTCCCTATGCGCCACTGTGT 494

RESULT 15
AUI40971
LOCUS 861 bp mRNA linear EST 05-AUG-2002
DEFINITION AUI40971 PLACE4 Homo sapiens cDNA clone PLACE4000583 5', mRNA
sequence.
ACCESSION AUI40971
VERSION AUI40971.1 GI:11002492

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 861)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isoqai,T.
TITLE HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isoqai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source Location/Qualifiers
1..861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACB4000583"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pME18SFL3"
ORIGIN
Query Match 13.0%; Score 231.2; DB 1; Length 861;
Best Local Similarity 79.5%; Pred. No. 1.3e-54;
Matches 307; Conservative 0; Mismatches 33; Indels 46; Gaps 1;
QY 254 CCCATAGCTGAGAGTGTGTTGATCATGCTGGGACTTCCTATGTGGTGGGAGAAACG 313
DB 158 CCCATAGCTGAGAGTGTGTTGATCATGCTGGGACTTCCTATGTGGTGGGAGAAACG 217
QY 314 -----TGGGAGAGGCGAGC 327
DB 218 TGGGAGAGGCCCTACCAAGCTGGATGATGTTAGTTGCTTGGGAGAGGCGAGC 277
QY 328 GGAACGATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT 387
DB 278 GGACGCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT 337
QY 388 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 447
DB 338 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCCAGTGCATCTGC 397
QY 448 ACAGGCAACGGCCGAGGAGTGGAGTGGAGGCGACACCTGTGCGAGACCATCG 507
DB 398 ACAGGCAACGGCCGAGGAGTGGAGTGGAGGCGACACCTGTGCGAGACCATCG 457
QY 508 AGCGGATCTGGGCCCTTCCCGATGTTGCTGGACCTGAGTGGCTGTAGACCGT 567
DB 458 AGCGGATCTGGGCCCTTCCCGATGTTGCTGGACCTGAGTGGCTGTAGACCGT 517
QY 568 CCATCTGTCAACCAACAGCCAAATTGGT 593
DB 518 CAGCCTCTCCATATGGCCACTGTGT 543

Search completed: February 1, 2006, 12:22:51
Job time : 7324.56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:47:34 ; Search time 123.96 Seconds
(without alignments)
12632.698 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134

Sequence: 1 tcgcttcacgttcgtcg.....ataacctaacgacaaataa 1782

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p model -DRV=xlp
-Q/cn2_1/USFT0.spool_p/US0940235/runat 27012006_144217_27548/app_query.fasta_1.7708
-DB=A_Geneseq -Qfmt=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CN1_1_846 @runat 27012006_144217_27548 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21.*

1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2168	69.2	795	2 AAY24797	Aay24797 Streptoki
2	2165	69.1	483	2 AAR12885	Aar12885 Factor xa
3	2158	68.9	414	2 AAR10194	Aar10194 Streptoki
4	2158	68.9	414	2 AAR63120	Aar63120 Streptoki
5	2158	68.9	414	2 AAY24794	Aay24794 Streptoco
6	2158	68.9	414	2 AAW94664	Aaw94664 Streptoco
7	2158	68.9	414	2 AAY01556	Aay01556 Native st
8	2158	68.9	414	3 AAY90282	Aay90282 S. equisi
9	2157	68.8	531	2 AAR11829	Aar11829 FB-FB-SK

10	2154	68.7	414	2 AAW94665	Aaw94665 Streptoco
11	2153	68.7	413	8 ADL92189	Adl92189 Streptoki
12	2150.5	68.6	1194	2 AAW21726	Aaw21726 Streptoki
13	2143	68.4	499	2 AAR12891	Aar12891 Streptoki
14	2140.5	68.3	813	2 AAW21728	Aaw21728 Wild type
15	2132	68.0	414	2 AAW86143	Aaw86143 Streptoki
16	2132	68.0	414	3 AAB01295	Aab01295 Wild type
17	2132	68.0	415	9 AEB93425	Aeb93425 Streptoco
18	2132	68.0	440	2 AAR12889	Aar12889 Streptoki
19	2132	68.0	483	2 AAR12522	Aar12522 Factor xa
20	2132	68.0	859	2 AAR12893	Aar12893 OmpAL str
21	2128	67.9	414	2 AAR20202	Aar20202 S.equisim
22	2121	67.7	414	2 AAW86144	Aaw86144 De-immuni
23	2121	67.7	414	3 AAB01296	Aab01296 Altered s
24	2120.5	67.7	415	3 AAY50870	Aay50870 Streptoco
25	2120.5	67.7	415	3 AAY99593	Aay99593 Streptoco
26	2120.5	67.7	1194	2 AAW21724	Aaw21724 Modified
27	2116	67.5	440	2 AAR06377	Aar06377 Streptoki
28	2114.5	67.5	1194	2 AAW21725	Aaw21725 Modified
29	2090.5	66.7	413	2 AAY25020	Aay25020 Streptoki
30	2090.5	66.7	413	5 ABB80012	Abb80012 Streptoki
31	2090.5	66.7	413	6 AEG74199	Aeg74199 Represent
32	2072	66.1	414	4 AAB74940	Aab74940 Mutant st
33	2071	66.1	800	2 AAW21723	Aaw21723 Plasminog
34	2071	66.1	1181	2 AAW21727	Aaw21727 Streptoki
35	2061	65.8	401	3 AAY84004	Aay84004 Amino aci
36	2061	65.8	413	3 AAY84005	Aay84005 Amino aci
37	2022	64.5	440	1 AAP50620	Aap50620 Sequence
38	2005.5	64.0	391	2 AAR10199	Aar10199 Streptoki
39	1938	61.8	440	2 AAR06378	Aar06378 Streptoki
40	1937	61.8	440	5 ABP29561	Abp29561 Streptoco
41	1937	61.8	440	8 ADR83961	Adr83961 S. pyogen
42	1924	61.4	372	2 AAR10200	Aar10200 Streptoki
43	1916	61.1	372	2 AAR10197	Aar10197 Streptoki
44	1907.5	60.9	371	2 AAR10195	Aar10195 Streptoki
45	1907	60.8	374	2 AAR10198	Aar10198 Streptoki

ALIGNMENTS

RESULT 1

AAY24797

ID AAY24797 standard; protein; 795 AA.

AC AAY24797;

XX XX

DT 26-AUG-1999 (first entry)

DE Streptokinase and maltose binding protein fusion protein.

XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;

KW nSK; rSK; bacterial; blood clot; thrombotic condition;

KW myocardial infarction; venous thrombosis; pulmonary embolism;

KW cerebral thrombosis; graft thrombosis; arterial thrombosis.

XX OS Streptococcus dysgalactiae subsp. equisimilis.

OS Synthetic.

XX WO9931247-A1.

XX PD 24-JUN-1999.

XX PF 15-DEC-1998; 98WO-US026694.

XX PR 15-DEC-1997; 97US-0069497P.

XX PA (HARD) HARVARD COLLEGE.

XX PI Reed GL;

XX DR WPI; 1999-395183/33.

XX DR N-PSDB; AAX80497.

PT N-terminally deleted streptokinase.
 XX Example; Page 48-51; 73pp; English.
 PS
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents a streptokinase and maltose binding protein fusion protein
 CC from an example of the present invention
 XX
 SQ Sequence 795 AA;

Alignment Scores:

Pred. No.: 1.45e-193 Length: 795
 Score: 2168.00 Matches: 421
 Percent Similarity: 97.91% Conservative: 1
 Best Local Similarity: 97.68% Mismatches: 5
 Query Match: 69.18% Indels: 4
 DB: 2 Gaps: 2

US-09-940-235-11 (1-1782) x AAY24797 (1-795)

QY 496 CAGACCATCGAGC-----GGATCTGGCCCTTCCCGATGTTTGGTATTGCTGGA 546
 DB 366 GlnThrAsnSerSerValProGlyArgGlySerIle-:::GluGlyArgIleAlaGly 384
 QY 547 CCTGAGTGCCTGAGCGCTCCATCTGTACACAGCCCAATTTGGTGTAGCGTTGCT 606
 DB 385 ProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValSerValAla 404
 QY 607 GGTACTGTTGAGGGACGAATCAAGACATTAGTCTTTAAATTTTGGAAATCGATCTAAACA 666
 DB 405 GlyThrValGluGlyThrAsnGlnAspIleSerLeuLeuPhePheGluLeuLeuLeuThr 424
 QY 667 TCACGACCTGCTCATGGAGGAAGACAGACGAGCTTAAGTCCAAATCAAAACCATTT 726
 DB 425 SerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSerLysProPhe 444
 QY 727 GCTACTGATAGTGGCGGATGTCACATAACTTGAGAAAGCTGACTTAAAGGCTATT 786
 DB 445 AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeuLysAlaIle 464
 QY 787 CAAGAACCAATTGATCGCTAACTCCACAGTAACTGAGAGCTACTTTGAGGTCAATTGATTTT 846
 DB 465 GlnGluGlnLeuLeuAlaAsnValHisSerAsnAspAspTyrPheGluValIleAspPhe 484
 QY 847 GCAAGCGATGCAACCATTAATCATCGAAACGCAAGGCTTACTTGTCTGCAACAAGATGTT 906
 DB 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAspLysAspGly 504
 QY 907 TCGGTAACTTCCGACCCCACTGTCAGAGATTTTTCGTAAGCGGACATGTCGCCGTT 966
 DB 505 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 524
 QY 967 AGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGCTGGAATATACT 1026

DB 525 ArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 544
 QY 1027 GTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACAGCTCTCAAGATACTAAG 1086
 DB 545 ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLys 564
 QY 1087 CTATTGAAAAACACTAGCTATTCGGTGCACACCATCACATCTCAAGAATTACTAGCTCAAGCA 1146
 DB 565 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeuAlaGlnAla 584
 QY 1147 CAAGCAATTTTAAACAAAACCCAGGCTTATACGATTTATGAACTGATCTCTCAATC 1206
 DB 585 GlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAspSerIle 604
 QY 1207 GTCACATCATGACATGACATTTCCGTACCATTTTACCAATGATGATCAAGAGTTTACTTAC 1266
 DB 605 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyr 624
 QY 1267 CGTGTAAAAAATCGGAACCAAGCTTATAGCATCAATAAAAAATCTGCTCGAATGAAGAA 1326
 DB 625 ArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGluGlu 644
 QY 1327 ATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGGGAAAGCCG 1386
 DB 645 IleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGlyGluLysPro 664
 QY 1387 TATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCCACCATCAATACGTTGATGTCAT 1446
 DB 665 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 684
 QY 1447 ACCAACCAATTGCTAAAAAGTGAAGAGCTCTTAAACAGTAGCGAACCGTAACCTTAGACTTC 1506
 DB 685 ThrAsnGluLeuLeuLysSerGluGlnLeuLeuLeuThrAlaSerGluArgAsnLeuAspPhe 704
 QY 1507 AGAGATTATPACGATCTCTGATAAGGCTAAAGTAACTACTCTACCAACATCTCGATGCTTTT 1566
 DB 705 ArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPhe 724
 QY 1567 GGTATTATGACATATACCTTAACCTGGAAGAGTAGAGATTAATCAGATGACCAACCCGT 1626
 DB 725 GlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAspThrAsnArg 744
 QY 1627 ATCATAAACCGTTTATATGGCAAGCGACCCGAGAGGAGAGATGCTAGCTATCATTTAGGCC 1686
 DB 745 IleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyrHisLeuAla 764
 QY 1687 TATGATAAGATCGTTTATACCGAAGAGAACGAGAGAGTTTACAGCTACCTCGCTTATACA 1746
 DB 765 TyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSerTyrLeuArgTyrThr 784
 QY 1747 GGGACACCTATACCTGATTAACCTTAACCGACAAA 1779
 DB 785 GlyThrProIleProAspAsnProAsnAspLys 795
 RESULT 2
 AAR12885
 ID AAR12885 standard; protein; 483 AA.
 XX
 AC AAR12885;
 XX
 XX 25-MAR-2003 (revised)
 DT 17-SEP-1991 (first entry)
 XX
 DE Factor Xa-cleavable hirudin-IEGR-streptokinase.
 XX
 XX Fusion protein; blood clotting; coagulation; fibrinolysis;
 KW antithrombotic; thrombolysis.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..65
 FT /label= hirudin HV-1

Db	45	-----ThrProLysProGlnSerHisAsnAspGlyAspPheGluGluIleProGluGlu	62
QY	535	-----CGPATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTC	576
Db	63	TyrLeuGlnIleGluGlyArgIleAlaGlyProGluTrpLeuLeuAspArgProSerVal	82
QY	577	AACAAACGCCAAATTGGTGTGTAGCGTGTGGTACTGTTGAGGGGAGCAATCAAGACATT	636
Db	83	AsnAsnSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIle	102
QY	637	AGTCCTTAAATTTTGGAAATCGATCTAACTACAGACTGCTGTATGAGGAAAGACAGAG	696
Db	103	SerLeuLysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGlu	122
QY	697	CAAGGCTTAAGTCCAAATCAAAACCAATTGGTACTGATGATGTCGCGGAGTCTCACATAAA	756
Db	123	GlnGlyLeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetProHisLys	142
QY	757	CTTGAGAAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAAACGTCACAGT	816
Db	143	LeuGluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSer	162
QY	817	AACGACGACTACTTTGAGGTCATTGATTTTGCAGAGCATGCAACCATTACTGATCGAAAC	876
Db	163	AsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsn	182
QY	877	GGCAAGTCTACTTTGCTGCAAGAAGTGGTTCGGTAACTTGGCGACCCCAACCTGCCAA	936
Db	183	GlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGln	202
QY	937	GAATTTTGTGAAGCGGACACTGTGCGGTTAGACCATATAAGAAAAACCAATACAAAAC	996
Db	203	GluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsn	222
QY	997	CAAGCGAAATCTGTTGATGTGGAAATATACGTGTACAGTTTACTCCCTTAAACCTCATGAC	1056
Db	223	GlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAsp	242
QY	1057	GATTTTCAGACGAGTCTCAAGATACTTAAGCTATTGAACAACTAGCTATCGGTGCACACC	1116
Db	243	AspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThr	262
QY	1117	ATCACATCTCAAGAAATTACTAGCTCAAGCAACAAGCAATTTTAAACAAAAACCAACGAGGC	1176
Db	263	IleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGly	282
QY	1177	TATACGATTTATGAACGTGACTCTCTCAATCGTCACTCATGCAATGACATTTTCCGTACG	1236
Db	283	TyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThr	302
QY	1237	ATTTTACCATGATCAAGAGTTTACTTACCGTGTAAATAATCGGAGCAAGCTTATAGG	1296
Db	303	IleLeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGlu	322
QY	1297	ATCAATAAAAAATCTGCTCTGAATGAAGAAATAAACAACATGCACTGTACTCTGAGAAA	1356
Db	323	IleAsnLysLysSerGlyLeuAsnGluLuleAsnAsnThrAspLeuIleSerGluLys	342
QY	1357	TATTACGTCCTTAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAAA	1416
Db	343	TyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLys	362
QY	1417	CTGTTCCACCATCAATACGTTGATGTGCTATACCACAGAAATGTCTAAAGATGAGCAGCTC	1476
Db	363	LeuPheThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeu	382
QY	1477	TTAAACAGCTAGCGAAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCT	1536
Db	383	LeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAla	402
QY	1537	AAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGAAAA	1596
Db	403	LysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLys	422

```
QY 1597 CTAGAGGATAATCAGCATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCC 1656
Db |||
423 ValGluAspAsnHisAspThrAsnArgIleIleThrValTyrMetGlyLysArgPro 442
QY 1657 GAGAGAGAGATGCTAGCTATCATTTAGCCTATGATAAGATCGTTATACCGAAGAGAA 1716
Db |||
443 GluGlyGluAsnAlaSerTyrHisLeuAlaTyrAspLysAspArgTyrThrGluGluGlu 462
QY 1717 CGAGAAGTTTACAGCTACCTGCTGTTATACAGGACACCTATACCTGATACCTACCGAC 1776
Db |||
463 ArgGluValTyrSerTyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAsp 482
QY 1777 AAA 1779
Db |||
483 Lys 483

RESULT 3
AAR10194
ID AAR10194 standard; protein; 414 AA.
XX AC AAR10194;
XX DT 28-MAR-1991 (first entry)
XX DE Streptokinase encoded by synthetic gene.
XX KW streptokinase; thrombolytic agent; myocardial infarction.
XX OS Synthetic.
XX PN EP407942-A.
XX PD 16-JAN-1991.
XX PF 11-JUL-1989; 89JP-00179432.
XX PR 11-JUL-1989; 89JP-00179432.
XX PR 27-NOV-1989; 89JP-00307937.
XX PR 11-APR-1990; 90JP-00096830.
XX PA (SAKA ) OTSUKA PHARM FACTOR.
XX PI Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
XX WPI; 1991-016179/03.
XX DR N-PSDB; AAQ10230.
XX ST Synthetic gene encoding streptokinase - scale, high purity prodn. of
PT streptokinase used as a thrombolytic agent.
PS Claim 1; Page 51; 76pp; English.
XX Streptokinase and its derivatives can be produced in large quantities
CC with high purity for use as thrombolytic agents in patients with lung
CC thrombus or myocardial infarction. See also AAR10195-R10200
XX SQ Sequence 414 AA;

Alignment Scores:
Pred. No.: 9,55e-193 Length: 414
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x AAR10194 (1-414)

QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTGGTGT 597
Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
```

```
QY 598 AGCGTTGCTGCTACTGTTGAGGGAGCAATCAAGACATTAGTCTTAATATTTTGAATC 657
Db |||
21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTAACATCAGCAGCTGCTCATGAGGAGAAAGACAGACGAGCTTAAGTCCAAATCA 717
Db |||
41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
QY 718 AAACCATTTCTACTGATAGTGGCGGATGTCATAACTTGAGAAAGCTGACTTACTA 777
Db |||
61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
QY 778 AAGCGTATTCAAGAACAAATTCATCGCTTAACGTCACAGTAAACGACGACTACTTTGAGGTC 837
Db |||
81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 838 ATTGATTTTCCAAGCGATGCAACATTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 897
Db |||
101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 898 AAGATGCTTCGGTAACTTGGCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 957
Db |||
121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 958 GTGCGGTTAGACCATATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1017
Db |||
141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATACCTGACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAGACGAGGTCCTCAA 1077
Db |||
161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 1137
Db |||
181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1138 GCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTATGAACGTGAC 1197
Db |||
201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 1198 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 1257
Db |||
221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAATCTGGTCTG 1317
Db |||
241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 1318 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1377
Db |||
261 AsnGluGluLeuAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1378 GAAAGCCGTPATGATCCCTTTTGATCGCAGTCACCTTGAACATGTTTACCACCAATAGCTT 1437
Db |||
281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1438 GATGTCGATACCAACGAATTCCTTAAAGTAGGAGAGCTCTTAACAGCTAGCAACGTAAC 1497
Db |||
301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGGTGATAAGGCTAAACTTACTCTACCAACATCTC 1557
Db |||
321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATACGATGAC 1617
Db |||
341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1618 ACCAACGGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAGATGCTAGCTAT 1677
Db |||
361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
QY 1678 CATTTAGCCTATGNTAAAGATCGTTATACCGAAGAACGACGAGAGTTTACAGCTACCTG 1737
```

|||||
Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGluGluArgGluValTyrSerTyrLeu 400
QY 1738 CGTTATACAGGACACCTATACCTGATAACCCCTAAACGACAAA 1779
Db 401 ArgTyrThrGlyThrProLeuProAspAsnProAsnAspLys 414

RESULT 4

AAR63120
ID AAR63120 standard; protein; 414 AA.
XX AAR63120;
XX AC
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 16-NOV-1994 (first entry)
XX Streptokinase.
XX
KW Streptokinase; SK; Streptococcus equisimilis; plasminogen;
KW myocardial infarction.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers
FH Region 1..352
FT /note= "claim 3, see CC"
FT Region 14..414
FT /note= "claim 1, see CC"
FT Region 120..352
FT /note= "claim 3, see CC"
FT Region 244..414
FT /note= "claim 3, see CC"
FT Region 244..352
FT /note= "claim 2, see CC"

XX WO9407992-A1.

XX PD 14-APR-1994.

XX PF 05-OCT-1993; 93WO-US009502.

XX PR 05-OCT-1992; 92US-00956692.

XX PR 29-SEP-1993; 93US-00128299.

XX PA (GEHO) GEN HOSPITAL CORP.

XX PA (HARD) HARVARD COLLEGE.

XX PI Reed GL;

XX PI WPI; 1994-135561/16.

XX DNA encoding a polypeptide which binds to plasminogen and corresponds to
PT region of streptokinase - useful to detect plasminogen in a sample and to
PT treat myocardial infarction.

XX Disclosure; Page 40-41; 62pp; English.

XX Nucleic acid comprising a sequence encoding amino acids 14-414 of
CC streptokinase, which binds to plasminogen and does not have a sequence
CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 414 AA;

Alignment Scores:

Pred. No.:	9,55e-193	Length:	414
Score:	2158.00	Matches:	414
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.86%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-11 (1-1782) x AAR63120 (1-414)
QY 538 ATTGCTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACACAGCCAAATTGGTGTCT 597
Db 1 IleAlaGlyProGluTrpLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTTGCTGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTTCAAAATC 657
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTAATCATCAGCACTGCTCATGGAGGAAAGACAGACGAAGGCTTAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 718 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAATCTGAGAAGCTGACTTACTA 777
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY 778 AAGCTATTCAAGAACCAATTGATCGCTAACTGCTCACAGTCCACAGTACGACGACTACTTTGAGTCT 837
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 838 ATTGATTTTGCAGCGATGCAACCACTTACTGATCGAAACGCCAAGGCTACTTTGCTGAC 897
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 898 AAAGATGGTTGGTAACTTGGCCGACCACTCTCCAGAAATTTTGGCTAAGCGGCAT 957
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 958 GTGGCGCTTAGACCATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 1017
Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGACAGGCTCTCAA 1077
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAATTACTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1138 GCTCAAGCACAAAGCATTTTAAACAAAAACACCCAGGCTATACGATTATGACGTCGAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTACCAATGATCAAGAG 1257
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1258 TTTACTTACCGTGTAAAAATCGGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 1318 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1378 GAAAGCCGTATGATCCCTTTTGTATCGCAGTCACCTTGAACCTGTTCCACCATCAAAATCGTT 1437
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1438 GATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCAGCACTAAC 1497
Db 301 AspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 320
QY 1498 TTAGACTTCAGAGATTTATAGTCTCTCGTGATAGGCTAAACTACTCTACAAACATCTC 1557
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1558 GATGCTTTTGGTATTATGGACTACTACCTTAACCTGGAAGTAGAGGATATACGATGAC 1617

Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1618 ACCAACCGTATCATAAACCGTTTATATGGCGAAGCGACCCGAGAGAGATGCTAGCTAT 1677
 Db 361 ThrAsnArgIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr 380
 QY 1678 CATTAGCCTATGATAAGATCGTTATACCGAAGAGAAAGAGAGAGTTTACAGCTACCTG 1737
 Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSerTyrLeu 400
 QY 1738 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAA 1779
 Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 5

AY24794

ID AY24794 standard; protein; 414 AA.

XX AY24794;

AC 17-OCT-2003 (revised)

DT 26-AUG-1999 (first entry)

XX Streptococcus equisimilis native streptokinase.

DE Streptococcus; streptokinase; fibrin-dependent plasminogen activator;

KW nSK; rSK; bacterial; blood clot; thrombotic condition;

KW myocardial infarction; venous thrombosis; pulmonary embolism;

KW cerebral thrombosis; graft thrombosis; arterial thrombosis.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS WO9931247-A1.

XX 24-JUN-1999.

PD 15-DEC-1998; 98WO-US026694.

PF 15-DEC-1997; 97US-0069497P.

PR (HARD) HARVARD COLLEGE.

XX Reed Glu;

XX WPI; 1999-395183/33.

XX N-PSDB; AAX80492.

XX N-terminally deleted streptokinase.

XX Claim 30; Page 60-61; 73pp; English.

CC The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmacological composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (I) encoding a modified bacterial streptokinase; (3) an expression vector comprising (I); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence represents native streptokinase (nSK). (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 414 AA;

Alignment Scores:

Pred. No.: 9.55e-193 Length: 414

Score: 2158.00 Matches: 414

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 68.86% Indels: 0

DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x AAY24794 (1-414)

QY 538 ATTGCTGGACCTCGAGTGGCTGTAGACCGTCCATCTCTCAACACAGCCCAATTGTTGTT 597
 Db 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 598 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATAATTTTGAATC 657
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 658 GATCTAACATCAGCAGCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGCTACATAAACTTGAGAAAGCTGCTTACTA 777
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 80
 QY 778 AAGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 837
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 838 ATTGATTTTGCAGAGCATGCAACCATCTACTGATCGAAACCGGAGGCTACTTTGCTGAC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 898 AAAGATGGTTCGGTAACTTGGCGACCCACCTGCTCCAGCAATTTTGTAAAGCGGACAT 957
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 QY 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 1017
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGACAGGCTCTCAAA 1077
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1078 GATACTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCATCTCAAGAACTACTA 1137
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
 QY 1138 GCTCAAGCACAAAGCATTTTAAACCAAAACCAACCCAGGCTATACGATTTATGAACGTGAC 1197
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 1198 TCCTCAATCGTCTCATGCAATGATCATTTTCCGTCAGATTTTACCAATGGATCAAGAG 1257
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1258 TTTACTACGGTGTAAAAATCCGGAAACAGCTTATAGGATCAATAAAAAATCTGCTCTG 1317
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 1318 AATGAAGAATAAACACACTGACCTGACCTCTCTGAGAAATATTACGTCCTTAAAGAGGG 1377
 Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1378 GAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAATAGCTT 1437
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300

QY 1438 GATGTCGATACCAACAATTGCTAAAGAGAGAGCTTTAAACGCTACGAAACGTAC 1497
 Db 301 AspValAspThrAsnGluLeuLeuLysSerGluLeuLeuThrAlaSerGluArgAsn 320
 QY 1498 TTAGACTTCAGAGATTATACGATCCTCGGATAGAGCTAACTACTACACAAATCTC 1557
 Db 321 LeuAspPheArgSerPheLeuThrAspProArgAspLysAlaLysLeuLeuThrAsnAsnLeu 340
 QY 1558 GATGCTTTGGTATTATGAGACTTATACCTTAACTGAAAGAGATAGAGATATTCAGATAC 1617
 Db 341 AspAlaPheGlyLysLeuLeuAspThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1618 ACCAACCGTATCATTAACCGTTATATGAGGACGACCCGAGAGAGAAATGCTAGCTAT 1677
 Db 361 ThrAsnArgGluLeuThrValTyrMetGlyLysArgProGluGluGluAsnAlaSerTyr 380
 QY 1678 CATTTAGCCTATATGAATGATGCTTATACCGAAGAGAGAGAAAGATTACAGCTACCTG 1737
 Db 381 HisLeuAlaLysThrAspLysAspArgTyrThrGluGluGluValArgGluValTyrSerTyrLeu 400
 QY 1738 CGTTATACAGGAGACCTTATACCTGATTAACCTTAAGACAA 1779
 Db 401 ArgTyrThrGlyThrProLeuProAspAsnProAsnAspLys 414
 RESULT 6
 AAW94664
 ID AAW94664 standard; protein; 414 AA.
 XX AAW94664;
 AC AAW94664;
 XX
 DT 17-OCT-2003 (revised)
 DT 04-MAY-1999 (first entry)
 DE Streptococcus equisimilis native streptokinase.
 XX
 XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KM plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KM serine protease; fibrin; blood clot; thrombolytic;
 KM vascular thromboembolytic symptom; acute myocardial infarction;
 KM fibrinolysis; resistance.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 14 /note="encoded by Acc"
 FT
 XX US5876999-A.
 PN 02-MAR-1999.
 PD 06-DEC-1995; 95US-00568393.
 PF 06-DEC-1995; 95US-00568393.
 PR 06-DEC-1995; 95US-00568393.
 XX (NASC-) NAT SCI COUNCIL.
 PA
 XX Mu H;
 PI
 PL MPI; 1999-189643/16.
 DR N-PSDB; AAX16632.
 XX
 XX Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT thrombolytic agent.
 FT
 XX Claim 1; Col 7-10; 17pp; English.
 XX
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence represents native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPig) to
 CC plasmin (HPIm), which is a serine protease able to catalyse the

CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 9.55e-193 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-11 (1-1782) x AAW94664 (1-414)
 QY 538 ATTGCTGAGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTGT 597
 Db 1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGluLeuVal 20
 QY 598 AGCGTTGCTGCTACTGTTGAGGGGAGCAATCAAGACTTAATTTTGAATC 657
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeuLysPheGluVal 40
 QY 658 GATCTAACATCACGACCTGTCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 717
 Db 41 AspleuThrSerArgProAlaHisGlyGlyThrGluGluGluLeuSerProLysSer 60
 QY 718 AAACATTTGCTACTGATAGTGGCGCATGTCAATTAACCTTGAGAAAGCTGACTTACTA 777
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluValAlaAspLeuLeu 80
 QY 778 AAGGCTATTCAAGAAACATGATGCTTACGTCACAGTAAAGAGACTACTTGGAGTTC 837
 Db 81 LysAlaIleGluGluGluGluLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluVal 100
 QY 838 ATTGATTTTTCAGACGATGCAACCATTAAGTATGATGAAACGGCAAGTCTACTTGTCTGAC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
 QY 898 AAGATGCTTTCGGTAACTTGGCGGACCCAACTGTCCAAGATTTTGTACGCGACAT 957
 Db 121 LysAspGlySerValThrLeuProThrGluProValGluGluPheLeuSerGlyHis 140
 QY 958 GTGGCGGTTAGACCATTAAGAAACCAATTCAAACCAAGGAAATCTGTGATGTG 1017
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1018 GAATATCTAGTACAGTTTACTCCTTAAACCTTGATGACGATTTCAACAGGCTCAAA 1077
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1078 GATACTAAGCTATTGAAAACACATGACTGCTGAGACCAATCATCATGCTCAAGATTTACTA 1137
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGluLeuLeu 200
 QY 1138 GCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTGAACGTCAC 1197
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 1198 TCCTCATGCTGACTCATGATGACATGATTTTCGTTACGATTTTCAAGATGATCAAGAG 1257
 Db 221 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1258 TTACTTACCGTGTAAATATCGGAAACAAGCTTATAGATCAATAAAATCTGCTGTG 1317
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaLysThrArgIleAsnLysLysSerGlyLeu 260
 QY 1318 AATGAAGAAATTAACAACACTGACCTGATCTTGAGAAATATTACGCTTTAAAGAGG 1377

Db 261 AenGIuGIuIIeAsnaAnthrAspleuIIeSerGIuIyTyTyrValIleuIySlyGIy 280
 QY 1378 GAAAGCCGTATGATCCCTTTGATGCGACTTGAACCTGTCACCATCAATCGTT 1437
 Db 281 GIuIySProIyAspProPhaSPheAspSerHISleuIyIleuPhetHrIleIySyrVal 300
 QY 1438 GATGTCGATACCAACGAATTGGCTTAATAAGTAGACAGCTCTTAACAGCTAGCGGAACGAAC 1497
 Db 301 AspValIAspThrAsnGIuIeuleuIySerGIuIeuleuThrIAserGIuIaGAsn 320
 QY 1498 TTAGACTTCAGAGATTATATACGATCTCGTATAGCTTAACCTACTTACACAAATCTC 1557
 Db 321 LeuAspPheArgAspIeuIyAspProAspIyAspIyIleuIyIleuIyIleuIyIleu 340
 QY 1558 GATGCTTTGGTATTATGACATATACCTTAACCTGAAAGTAGAGATATACGATGAC 1617
 Db 341 AspIAspIeGIyIleuAspIyTrThrIeuthrGIyValGIuIAspAsnHISAspAsp 360
 QY 1618 ACCAAGCGTATACGATACCGTTATATAGGAGCGGACCGGAGAGAGAGATGCTACTAT 1677
 Db 361 ThrAsnArgIleIleThrValIyIleuGIyIyAspArgIuGIuIAsnIAserTyr 380
 QY 1678 CATTAGCCTATGATTAAGATCGTTATACCGAAGAGAAAGAAAGATTACAGCTACTG 1737
 Db 381 HisIeuIaIyAspIyAspAspArgIyTrThrGIuGIuIAspGIuIValIyIleuIyIleu 400
 QY 1738 CGTTATACAGGAGACCTATACCTGATTAACCTTAACGACAA 1779
 Db 401 ArgTyrThrGIyThrProIleProAspAsnProAsnAspIyS 414
 RESULT 7
 AAY01556
 ID AAY01556 standard; peptide; 414 AA.
 AC AAY01556;
 DT 17-OCT-2003 (revised)
 DT 18-JUN-1999 (first entry)
 XX Native streptokinase protein sequence.
 DE Antigenic peptide: streptokinase; streptokinase-specific antibody;
 KM thrombolytic activity; thrombolytic therapy; glomerulonephritis;
 KM rheumatic fever.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX MO9908698-AI.
 PN 25-FEB-1999.
 PD 18-AUG-1998; 98WO-US017114.
 PF 18-AUG-1997; 97US-0055911P.
 PR 18-AUG-1997; 97US-0055911P.
 PA (HARD) HARVARD COLLEGE.
 PA (GENO) GEN HOSPITAL CORP.
 PI Reed GL, ParhamI-Seren B;
 DR WPI; 1999-190113/16.
 XX New polypeptides which bind streptokinase-specific antibodies - useful in
 PT thrombolytic therapy.
 XX Disclosure: Page 12; 4app; English.
 CC The present sequence represents a native streptokinase. The specification
 CC describes a polypeptide which binds to a streptokinase-specific antibody
 CC and prevents the antibody binding to native streptokinase. The
 CC specification also describes a synthetic polypeptide (PI) comprising an
 CC epitope which binds to an streptokinase-specific antibody and reduces
 CC thrombolytic activity of streptokinase. PI is used in thrombolytic

CC therapy, and to prevent or treat glomerulonephritis and rheumatic fever.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX SQ Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 9.55e-193 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-11 (1-1782) x AAY01556 (1-414)
 QY 538 ATTGCTGACCTGATGCTGCTGATGACCGTGCATCTGTCAACAACAGCAATGGTTGT 597
 Db 1 ILeAlaGIyProGIuIurIleuIeuleuAspArgProSerValIeAsnSerGIuIeValI 20
 QY 598 AGCGTTGCTGGTACTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 657
 Db 21 SerValAlaGIyThrValGIuGIyThrAsnGIaAspIleSerIeuIySPhePheGIuIle 40
 QY 658 GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 717
 Db 41 AspleuIuIhSerArgProAlaHISGIyGIyIyThrGIuIuGIuIyIleuSerProIySer 60
 QY 718 AAACATTGGCTACTGATAGTGGCGGATGTCACATTAACCTTGAGAAAGCTGACTTA 777
 Db 61 IySPheIuIhSerArgProAlaHISGIyGIyIyThrGIuIuGIuIyIleuSerProIySer 80
 QY 778 AAGGCTATTCAAGAACATTTGATGCTAACGTCACAGTACGACGATCTTGAAGTC 837
 Db 81 IySalAlaIleGIuIuGIuIeuleuIleAlaAsnValHISerAsnAspIyTrPheGIuIVal 100
 QY 838 ATTGATTTTGGCAAGCATGCAACCATTAATCGATGCGAAAGCGGAGGCTACTTGGCGAC 897
 Db 101 ILeAspPheAlaSerAspAlaThrIleThrAspArgAsnIyIyValIyTrPheAlaAsp 120
 QY 898 AAAGATGTTCCGTTAACCCTTGCACCAACCTGTCACAAAGATTTTGGTAAACGCAAT 957
 Db 121 IyAspGIySerValIleuIeuleuProIuIhIleuIeuleuIleuIeuleuSerGIyHIS 140
 QY 958 GTGCGCGTTAGACCATATTAAGAAACCAATACCAACAGCGAAATCTGTGATGTG 1017
 Db 141 ValArgValArgProIyIyGIuIySProIleGIaAsnGIaIuIySerValAspVal 160
 QY 1018 GAATATACGTGACAGTTTACTCCCTTAACCCCTGATGACGATTTTCAGACCAAGGCTCAAA 1077
 Db 161 GIuIyTrThrValGIuIhIleuIeuleuProIeuleuAsnProAspAspPheArgProGIyIleuIyS 180
 QY 1078 GATACCTAAGCTATTGAAAACACTAGCTATCGGTGACACCTACATCATCTCAAGATTACTA 1137
 Db 181 AspThrIyIleuIeuleuIySerThrIleuAlaIleGIyAspThrIleThrSerGIuIeuleu 200
 QY 1138 GCTCAAGCAACAAAGCAATTTTAAACAAAACCAACCAAGGCTATACGATTATGAACGTGAC 1197
 Db 201 AlaGIaIleGIuIeuleuIleuIeuleuIySerIleuIeuleuIySerIleuIeuleuIySerIleu 220
 QY 1198 TCCTCAATGCTACTGATGACATGACATGATTTTCCGATCCGATTTTTCAGAAAGATCAAGAG 1257
 Db 221 SerSerIleValIyThrHISAspAsnAspIlePheArgThrIleuIeuleuIySerIleuIeuleu 240
 QY 1258 TTACTTACCGTGTAAATAACGGGAACAGCTTATAGATCAATATAAATCATCGTGTG 1317
 Db 241 PheThrTrArgValIyIyAsnAspArgGIuIuIaIyTrArgIleAsnIyIySerGIyIleu 260
 QY 1318 AATGAAGAAATAAACAACATGACCTGATCTGAGAAATATTAACGTCCTTAATAAAGGG 1377
 Db 261 AenGIuGIuIIeAsnAnthrAspleuIIeSerGIuIySyrValIleuIySlyGIy 280
 QY 1378 GAAAGCCGTATGATCCCTTTGATGCGACTTGAACCTGTCACCATCAATCGTT 1437

Db 281 GluProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleValSerVal 300
 Qy 1438 GATGTCGATACCAACCAATGCTAAAGAGAGAGCTTTACAGCTAGCAAGCTAAC 1497
 Db 301 AspValAspThrAsnIleuLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 320
 Qy 1498 TTAGACTTCAGAGATTTATACGATCTCTGATAGAGGCTAAACTACTCTCAACAATCTC 1557
 Db 321 LeuAspPheArgSerLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 Qy 1558 GATGCTTTGGTATTATGACCTATACCTTACCTGAGAAAGATAGAGATATACAGATGAC 1617
 Db 341 AspAlaPheGlyIleLeuMetAspTyrThrLeuThrGlyValGluAspAsnHisAspAsp 360
 Qy 1618 ACCAACCGTATCATACCGTTTATATGGCAGAGACCGGAGAGAGAGAAATGCTAGCTAT 1677
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluValAsnAlaSerTyr 380
 Qy 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAGAGAAAGTTTACAGCTACTG 1737
 Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 400
 Qy 1738 CATTATACGAGGACACTTACTGATTAACCTTACAGCAAA 1779
 Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414
 RESULT 8
 AAY90282
 ID AAY90282 strand; protein; 414 AA.
 XX AAY90282;
 AC
 XX
 DT 12-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE S. equisimilis streptokinase.
 XX
 KM Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KM plasminogen; human; fibrinectin; thrombolytic therapy;
 KM cardiovascular disorder; fibrinectin.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX
 PN EP1024192-A2.
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RBS.
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V,
 PI Yadav M;
 XX
 DR WPI: 2000-516032/47.
 DR N-PsDB; AAA37633.
 XX
 PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibronectin.
 XX
 XX Example 3; Fig 3; 58pp; English.
 PS
 CC This sequence represents the human Streptococcus equisimilis
 CC streptokinase protein sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibronectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a

CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
 CC
 XX
 SO Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 9,55e-193 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 3 Gaps: 0
 US-09-940-235-11 (1-1782) x AAY90282 (1-414)
 Qy 538 ATTGCTGACCTGAGTGCGCTGCTAGCCGTCATCTGTGCAACAAACGCAATGGTGT 597
 Db 1 IleAlaGlyProGluThrLeuLeuAspArgProSerValaAsnSerGlnLeuValVal 20
 Qy 598 AGCGTTCGCTGACTGTTGAGGGGACGAATCAAGCAATTACTCTTAATTTTGAATC 657
 Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 Qy 658 GATTAACATCACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 717
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
 Qy 718 AAGCATTTGCTACTGATAGTGGCGGATGCATCAATAACTTGAAGAACTGACTTA 777
 Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluValaAspLeuLeu 80
 Qy 778 AAGCTATTCAAGAACAATTGATGCTTAAGCTCAACAGTAAACGACTACTTTGAGTGC 837
 Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspArgPheGluVal 100
 Qy 838 ATTGATTTTCAAGCGATGCAACCATTAAGTATGATGCAACGCGAAGTCTACTTGTAGTGC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 Qy 898 AAGATGTTTCGATTAACCTTCCGACCAACCTGTCAGAAATTTTGTAAAGCGGACAT 957
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
 Qy 958 GTGGCGCTTGAACCATTAATAAGAAAAACAATACAAAACCAAGGAAATCTGTGATGTG 1017
 Db 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValaAspVal 160
 Qy 1018 GAATATATCTGTACAGTTACTCCCTTAAACCGTAGACAGATTTTCAACACGAGCTCAAA 1077
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 Qy 1078 GATACTAAGCTATTTGAACCACTAGTATGCGTGACACCATCATCATCAAGATTACTA 1137
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
 Qy 1138 GCTCAAGCAACAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTTGAACGTGAC 1197
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluLysArg 220
 Qy 1198 TCCTCAATCTCTCACTAGTACATGACATGATTTTCCGTGATTTTACCAAGATCAAGAG 1257


```

Db      221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspGlnu 240
Qy      1258 TTTACTTACCGGTAAATAATGGGAACAGCTTATAGATCAATATAAATCTGCTCG 1317
Db      241 PheThrTYrAspValLysAsnAspGlnuGlnAlaTYrArgIleAsnLysSerGlyLeu 260
Qy      1318 AATGAAGAAATTAACAACACTGACCTGATCTCTGAGAAATATATACCTCTTAAATAAGG 1377
Db      261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTYrTYrValLeuLysGly 280
Qy      1378 GAAAACCGTATGATCCCTTATGTCGACACTTGAAACTGTTCACCATCAATACGTT 1437
Db      281 GlnLysProTYrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTYrVal 300
Qy      1438 GATGTCGATACCAAGAAATTTGCTAAATAATGTCGACGCTCTTAAACACTGACGCAACTAAC 1497
Db      301 AspValAspThrAsnGlnLeuLysSerGlnGlnLeuThrIleAspGlnArgAsn 320
Qy      1498 TTAGACTTCAGAGATTTATACGATCCCTGATTAAGGCTAAACTATCTTACCAACATCTTC 1557
Db      321 LeuAspPheAspLeuTYrAspProAspAspLysAlaLysLeuLysLeuLysAsnLeu 340
Qy      1558 GATGCTTTGTATTATGACTATACCTTAACTGAGAAAAGTAGAGATATACGATGAC 1617
Db      341 AspAlaPheGlyIleMetAspTYrThrIleuThrGlyValGlnAspAsnHisAspAsp 360
Qy      1618 ACCAACCGTATCATTAACGTTTATATGGCAAGCCGCAAGAGAGAAATGCTTACTTAT 1677
Db      361 ThrAsnArgIleIleThrValTYrMetGlyLysArgProGlnGlyGlnAsnAlaSerTYr 380
Qy      1678 CATTAGGCTATGATTAAGATGTTATATACGGAAGAAAGAGAACTTTTACAGTACCTG 1737
Db      381 HisLeuAlaTYrAspLysAspArgTYrThrGlnGlnGlnArgGlnValTYrSerTYrLeu 400
Qy      1738 CGTTATACAGGAGACCTATACCTGATTAACCTTAAACGACAAA 1779
Db      401 ArgTYrThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 9
AAR11829
ID      AAR11829 standard; protein, 531 AA.
XX      AAR11829;
AC      AAR11829;
DT      08-JUL-1991 (first entry)
DE      FB-FB-SK fusion conjugate.
XX      Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KM      streptokinase; fusion protein.
XX      Staphylococcus aureus.
XX      Key      Location/Qualifiers
FT      Peptide      2..61
FT      Peptide      /label= FB monomer
FT      Peptide      62..119
FT      Peptide      /label= FB monomer
FT      Peptide      120..531
FT      Peptide      /label= streptokinase
XX      US5011686-A.
XX      30-APR-1991.
XX      15-NOV-1989; 89US-00437769.
XX      21-SEP-1987; 87US-00099242.
XX      (CREA-) CREATIVE BIOMOLEC.
XX      Pang RHL;
XX      PI

```

```

DR      WPI, 1991-140198/19.
DR      N-PSDB; AAQ11651.
XX      PT      Imparting injectable fibrinolytic agent - with affinity for intravascular
XX      thrombus, by linking agent to fibrin binding domain.
XX      PT      Disclosure; Fig 5; 18pp; English.
XX      XX      The conjugate comprises an FB-FB dimer linked to streptokinase The FB
CC      fragment has selective affinity for fibrin, low affinity for fibrinogen,
CC      and minimal immunogenicity, imparting thrombus- targeting capability.
CC      See also AAR11821 and AAR11828
XX      XX      Sequence 531 AA:
SQ      Sequence 531 AA:
Alignment Scores:
Pred. No.: 1,31e-192 Length: 531
Score: 2157.00 Matches: 413
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.52% Mismatches: 0
Query Match: 68.83% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-11 (1-1782) x AAR11829 (1-531)
Qy      535 CGTATTCGCGACCTGAGTGGCTGCTAGACCCGTCATCTGTCACACAGCAATTGGTT 594
Db      117 LysIleAlaGlyProGlnIleuLysAspArgProSerValAsnAsnThrGlnLeuVal 136
Qy      595 GTTAGCGTCTGCTGTACTGTGAGGAGCAATCAAGACATTAAGTCTTAATTTTGA 654
Db      137 ValSerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGln 156
Qy      655 ATGATCTTAACATCACAGCCCTGCTATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAA 714
Db      157 IleAspLeuThrSerArgProAlaHisGlyLysThrGlnGlnGlnLysLeuSerProLys 176
Qy      715 TCAAAACCATTTCTCTACTGATAGTGGCGGATGTCATATACTTGAGAAAGCTGACTTA 774
Db      177 SerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeu 196
Qy      775 CTAAGGCTATTTCAGAAACAATTGATCGCTAACGTCACAGTAAAGCACTACTTGGAG 834
Db      197 LeuLysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTYrPheGln 216
Qy      835 GTCAATGATTTTTCAGACGATGCAACCATTAATGATGAAACGCGACGATCTTGGCT 894
Db      217 ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTYrPheAla 236
Qy      895 GACAAAGATGCTTGGTAACTTGGCCGACCCAACTGTCCAAAGATTTTGTAAAGCGGA 954
Db      237 AspLysAspGlySerValThrLeuProThrGlnProValGlnGlnLysPheLeuSerGly 256
Qy      955 GATGCGCGGTTAGACCATTAAGAAACCAATCAAAACCAAGGAAATCGTTGAT 1014
Db      257 HisValArgValArgProTYrLysGlnLysProIleGlnAsnGlnAlaLysSerValAsp 276
Qy      1015 GTGGAATATACGTATACGTTACTCCCTTAAACCTGATGACGATTTCAAGACAGCTTC 1074
Db      277 ValGlnTYrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu 296
Qy      1075 AAGATACTAGCTATTTGAACACATGACTATGCTGATGACACCATCATCTCAAGAAATTA 1134
Db      297 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 316
Qy      1135 CTAGCTCAAGCAACAAGCATTTTAAACAACCAACCCAGGCTATAGATTTTGAAGCT 1194
Db      317 LeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTYrThrIleTYrGlnArg 336
Qy      1195 GACTCCCAATCGTCACTGACATGACAAATGATTTTCGTTACGATTTTACCAATGATCAA 1254
Db      337 AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspGln 356

```

QY 1255 GAGTTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATGCAATATAAAATCTGGT 1314
Db |||||||
Db 357 GIupheThrTyzrArgValIysAsnArgGluGlnAlaItyzrArgIleAsnLysSergIy 376
QY 1315 CTGAATGAAGAATAAACAACGACGCTGATCTCGAGAAATATTACCGCTTAAAAAA 1374
Db |||||||
Db 377 LeuAsnGluGluIleAsnAsnThrAspLeuIleSeriLysTyzrValIleuLysLys 396
QY 1375 GGGGAAAGACCGGTATGATCCCTTGAATCGACGATCACTTGAACGTTCACATCAATAC 1434
Db 397 GIyGIuLysProIytrAspProPheAspArgSerHisleuLysleuPheThrIleLysTyzr 416
QY 1435 GTTGAATGATGATCAACAAGCAATGCTTAAAAAGTGAAGCACTTAAACGCTAAGCGAAGT 1494
Db 417 ValAspValAspThrAsnGluLeuLeuLysSergIuGlnleuLeuThrAlaSergIuArg 436
QY 1495 AACTTAGACTGAGATTTATACGATCCCTCGATGAAGGCTAACTACTCTACAAACAT 1554
Db |||||||
Db 437 AsnLeuAspPheArgAspLeuTyzrAspProArgAspLysAlaLysleuLeuTyzrAsnAsn 456
QY 1555 CTGATAGCTTTGGTATATAGACTATACCTTAACGTGAAGAAAGTAAAGATATACAGAT 1614
Db 457 LeuAspAlaPheGlyIleMetAspTyzrThrLeuThrGlyLysValGluAspAsnHisAsp 476
QY 1615 GACACCAACCGTATCATTAACCGTTTATATGGGCAAGCAAGCCGAAAGAGAAATGCTAGC 1674
Db 477 AspThrAsnArgIleIleThrValTyzrMetGlyLysArgProGluGluAsnAlaSer 496
QY 1675 TATCATTTAGCTTATATATAAGTCTTATACCGAAGAAAGCAAGAAAGTTTACAGCTAC 1734
Db 497 TyzrHisleuAlaItyzrAspLysAspArgTyzrThrGluGluGluArgIuValTyzrSeryr 516
QY 1735 CTGGTTATATACGGGACACCTATACCTGAATACCTTAACGACAA 1779
Db 517 LeuArgTyzrThrGlyThrProIleProAspAsnProAsnAspLys 531
RESULT 10
AAM94665
ID AAM94665 standard; protein; 414 AA.
AC AAM94665;
DT 04-MAY-1999 (first entry)
XX Streptococcus equisimilis mutant streptokinase K59E.
XX Streptococcus equisimilis mutant streptokinase K59E.
XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
XX Key Location/Qualifiers
FT Misc-difference 14 /note="encoded by ACC"
FT
XX US5876999-A.
XX 02-MAR-1999.
XX 06-DEC-1995; 95US-00568393.
XX 06-DEC-1995; 95US-00568393.
XX (MASC-) NAT SCI COUNCIL.
XX Mu H;
XX WPI; 1999-189643/16.
XX N-PSDB; AAX16633.

XX Mutant streptokinase polypeptide - useful as plasmin-resistant
PT thrombolytic agent.
PT
XX
XX Claim 4; Col 11-14; 17pp; English.
XX
XX The present invention describes a mutant streptokinase (SK) polypeptide
CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
CC segment of the corresponding native SK is replaced by another amino acid.
CC The present sequence is mutant SK K59E. SK is a secretory protein of
CC haemolytic Streptococcus able to activate human plasminogen (HPIg) to
CC plasmin (HPIm), which is a serine protease able to catalyse the
CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
CC agent in the treatment of vascular thromboembolytic symptoms such as
CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
CC is more resistant to degradation by human plasmin and is more effective
CC both in acting as a fibrinolytic agent and in activating human plasminogen
CC
SQ Sequence 414 AA;
Alignment Scores:
Pred. No.: 2.26e-192 Length: 414
Score: 2154.00 Matches: 413
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.76% Mismatches: 0
Query Match: 68.73% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-11 (1-1782) x AAM94665 (1-414)
QY 538 ATTGCTGACCTGAGTGGCTGCTGACCGCTCATCTGTCACAAACAGCAATGGTGT 597
Db 1 IleAlaIleProGluThrLeuLeuAspArgProSerValAsnAsnSergIleuValAl 20
QY 598 AGCGTTGCTGTATCTGTGAGGGGACCAATCAAGACATTACTTAAATTTTGAATC 657
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSeriLeuLysPheGluIle 40
QY 658 GATCTAACATCAGACCTGCTCATGAGGAAAGCAAGCAAGGCTTAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyLysThrGluGlnGlyLeuSerProGluSer 60
QY 718 AAACATTTGCTACTGATAGTGGGCGGATGCACATTAACCTTGAAGAAAGCTTACTTA 777
Db 61 LysProPheAlaThrAspSergIyAlaMetSeriHisLysLeuGlyValAspLeuLeu 80
QY 778 AAGGCTATTCAAGAACATTTGATCGTACCTCAAGTACGACGACTATTGAGGTC 837
Db 81 LysAlaIleGluGlnGluGlnleuIleAlaAsnValHisSerAsnAspArgTyzrPheGluVal 100
QY 838 ATTGATTTGCAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGTCTACTTGTGAC 897
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyzrPheAlaAsp 120
QY 898 AAAGATGTTCCGGTAACTTGGCCGACCAACCTGTCCAAGAAATTTTGTAAAGGACAT 957
Db 121 LysAspLysSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSergIyHis 140
QY 958 GTGCGGCTTAGACATTAAGAAAGAAACCAATACAAAGCAAGCAATGCTTGAAGTG 1017
Db 141 ValArgValArgProTyzrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATATCTGTACGATTACTCCCTTAAACCTTGATGACGATTTCAAGCAGGCTCAAA 1077
Db 161 GluTyzrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 1078 GATACTAAGCTATTGAAACACATGACTATGGGTGAACCATCACTTCAAGAAATTACTA 1137
Db 181 AspThrLysleuLeuTyzrThrLeuAlaIleGlyAspThrIleThrSergIuGluLeuLeu 200
QY 1138 GCTCAAGCAACAAGCATTTTAAACAAAGCAAGCCAGGCTATGATTTATGAAACGAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyzrThrIleTyzrGluArgAsp 220

QY 1198 TCTCAATCGTCACTCATGACATGATTTTCCGTACATTTTACCAATGATCAAGAG 1257
 DB 221 SerSerIleValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGlu 240
 QY 1258 TTACTACTCCGTGTAAATAATCGGAAACAGCTTATAGATCAATAAAAATCTGGCTG 1317
 DB 241 PheThrTyArgValIleAsnArgGlnGlnAlaTyArgIleAsnLysSerGlyLeu 260
 QY 1318 AATGAAGAAATAAACAACCTGATCTCTGAGAAATATTACCTCTTAAAAAGG 1377
 DB 261 AsnGlnGluIleAsnAsnThrAspLeuIleSerGlnLysTyArgIleValLysLysGly 280
 QY 1378 GAAAGCCGTATGATCCCTTGTGACGATGCTGAACTGTCACATCAATGCTG 1437
 DB 281 GluLysProTyArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTyArg 300
 QY 1438 GATGTGATACCAACGAATGCTTAAAGAGAGACAGCTTAAACAGCTAGCAAGCTAAC 1497
 DB 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 320
 QY 1498 TTACACTTCAGAGATTTTATGATCTCTGCTGATAGGCTTAACTACTTACAACTCTC 1557
 DB 321 LeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLeuTyArgAsnLeu 340
 QY 1558 GATGCTTTGGTATATGAGACTATACCTTACCTGAGAAAGTAGAGATATACAGATGAC 1617
 DB 341 AspAlaPheGlyIleMetAspProTyArgThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
 QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGAGAGATGCTAGCTAT 1677
 DB 361 ThrAsnArgIleIleThrValTyMetGlyLysArgProGlnGlnLysAsnAlaSerTyArg 380
 QY 1678 CATTTAGCTTATGATTAAGATGTTATACCGAAGAAAGAGATTATACAGCTACCTG 1737
 DB 381 HisLeuAlaTyArgLysAspArgTyArgThrGlnGlnGlnLysValTyArgLysTyArg 400
 QY 1738 CGTTATACAGGAGACCTATACCTGATTAACCTTAACGACAAA 1779
 DB 401 ArgTyArgThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 11
 ADL92189
 ID ADL92189 standard; protein; 413 AA.
 AC ADL92189;
 XX 20-MAY-2004 (first entry)
 DT 20-MAY-2004 (first entry)
 DE Streptokinase protein sequence.
 XX
 KW harvesting; recombinant; host cell; N-terminal leader peptide;
 KW pre-peptide; antibiotic; post-translational modification;
 KW pharmaceuticals; vaccine; immunogenic.
 OS Unidentified.
 XX
 PN WO2003099862-A1.
 PD 04-DEC-2003.
 PF 26-MAY-2003; 2003WO-NL000389.
 XX
 PR 24-MAY-2002; 2002EP-00077060.
 PR 07-FEB-2003; 2003US-00360101.
 XX
 PA (NANO-) APPLIED NANOSYSTEMS BV.
 PI Mol1 GN, Leenhouts CJ, Kuipers OP, Driessen AJM;
 XX WPI; 2004-042770/04.
 XX
 PT Harvesting a desired polypeptide produced by a recombinant host cell, for

PT producing pharmaceuticals, comprises selecting a recombinant nucleic acid
 PT comprising nucleic acid fragments encoding a leader peptide and the
 PT polypeptide.
 XX
 XX Claim 4; Page 82-83; 10pp; English.
 XX
 CC The invention relates to a novel method for harvesting a (poly)peptide
 CC produced by a recombinant host cell. The novel method involves selecting
 CC a cell comprising a first nucleic acid encoding a leader peptide and a
 CC second nucleic acid fragment encoding the desired (poly)peptide. The
 CC first and second fragments are within the same open reading frame of the
 CC first nucleic acid and the leader peptide is functionally equivalent to
 CC an N-terminal leader peptide found with the pre-peptide of a antibiotic.
 CC The host cells and nucleic acids are useful for producing, harvesting and
 CC post-translational modification of polypeptides. The polypeptides may be
 CC used in the production of pharmaceuticals, e.g. as antigen for vaccine or
 CC immunogenic composition. This sequence represents a polypeptide relating
 CC to the novel method of the invention.
 XX
 SQ Sequence 413 AA;
 Alignment Scores:
 Pred. No.: 2,8e-192 Length: 413
 Score: 2153.00 Matches: 413
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.70% Indels: 0
 DB: 8 Gaps: 0

US-09-940-235-11 (1-1782) x ADL92189 (1-413)

QY 538 ATGCGTCGACCGTGAAGGCGTCTGACACCGTCACTGTCACAAACAGCAATGCTGTT 597
 DB 1 IleAlaGlyProGlnTyArgLeuLeuAspArgProSerAlaAsnSerGlnLeuVal 20
 QY 598 AGCGTTCGCTGCTGCTGAGGAGCAATCAAGACATTAATGCTTAAATTTTGAATC 657
 DB 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPheGlnIle 40
 QY 658 GATCTAACATCAGACCTGCTCAATGAGAGAAAGAGAGAGAGCTTAAGTCCAAATCA 717
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnLysSerProLysSer 60
 QY 718 AAACATTTGCTACTGATAGTGGCGGATGTCATTAATTTGAGAAAGCTGACTTACTA 777
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeu 80
 QY 778 AAGGCTATTCAAGAAATGATGCGTCAACGTCACAGTAAAGCACTGTTAGAGTC 837
 DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspSerTyArgLeuVal 100
 QY 838 ATTTGATTTTGAAGCAGATGCAACCATTAATGATGCAAAACGCAAGCTCTACTTTGCTGAC 897
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnLysValTyArgPheAlaAsp 120
 QY 898 AAAAGATGTTCCGTTAACTTCCGACCCCAACTGTCACAAATTTTGGTAAAGCGCAT 957
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 958 GTGCGGTTTAAACCATTAAGAAACCAATCAAAACCAAGGAAATGCTTGAAGTG 1017
 DB 141 ValArgValArgProTyArgLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1018 GATATATGTCACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAGGCTTCAAA 1077
 DB 161 GluTyThrValGlnThrThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1078 GATCTTAAGCTATTGAAAACACTAGCTATCGTGAACCATCATCATCTGAAGATTACTA 1137
 DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAAACCAACCGGCTATGATTTTGAACGAGAC 1197

Db 201 AlaGlnAlaGlnSerIleLeuAenLysAenHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 1198 TCCTCAATCGTCACTGATGACATGATTTTCCTGATGATTTTACCAATGATCAAGAG 1257
 Db 221 SerSerIleValThrHisAspAenAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1258 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCTGGTCTG 1317
 Db 241 PheThrTyrArgValLysAenArgGluGlnAlaTyrArgIleAenLysSerGlyLeu 260
 QY 1318 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATATTCGCTTAAAAAGGG 1377
 Db 261 AsnGluGluIleAenAsnThrAspLeuIleSerGluLysTyrTyrValIleLysGly 280
 QY 1378 GAAAGCCGATATGATCCCTTGAATCGAGTCACTTGAACGCTTCAACATCAATAGCTT 1437
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1438 GATGTCGATACCAACGAAATGCTAAAAAGTAGACAGCTTTAAACAGCTAGGAACTTAC 1497
 Db 301 AspValAspThrAsnGluLeuLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 320
 QY 1498 TTRGACTTCAGAGATTTATTCGATCCTCGTATAGGCTTAACTCTTACACATCTC 1557
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1558 GATGCTTTGGTATTTAGACTATACCTTAACTGGAATAAGAGATTAATCCAGTAC 1617
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1618 ACCAACCGTATCAATAAACGTTTATATGCGCAACGCAACCGAAGAGAGAAATGCTAGCTAT 1677
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProIleGluIleAsnAlaSerTyr 380
 QY 1678 CATTTAGCTATGATTAAGAATGCTTATACCGAAGAAGAAAGAGAAAGTTTACAGTACTG 1737
 Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSerTyrLeu 400
 QY 1738 CCTTATACGAGGACACCTTATACCTGATTAACCTTAAAGAC 1776
 Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAsp 413
 RESULT 12
 AAW21726
 ID AAW21726 standard; protein; 1194 AA.
 AC AAW21726;
 XX
 DT 17-OCT-2003 (revised)
 DT 01-OCT-1997 (first entry)
 XX
 DE Streptokinase/maltose binding protein fusion protein, rSK.
 XX
 KW Plasminogen-binding fragment; streptokinase; degradation; MBP;
 KW thrombolytic agent; blood clot; bolus; maltose-binding protein.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1. 381
 FT Protein /label= Maltose binding protein
 FT /note= "acts as blocking group"
 FT 382. 1194
 FT /label= Streptokinase
 PN MO9641883-A1.
 XX
 PD 27-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US009640.
 XX
 PR 09-JUN-1995; 95US-00488940.
 XX

PA (HARD) HARVARD COLLEGE.
 XX
 PI Read GL;
 XX
 DR WPI; 1997-065469/06.
 XX
 PT Modified forms of streptokinase resistant to enzymatic cleavage - useful
 PT as thrombolytic agents in treating thrombosis and in medical equipment.
 XX
 PS Example 1, Page 11-12; 65pp; English.
 XX
 CC This sequence represents a fusion protein between maltose-binding protein
 CC and the plasminogen-binding fragment of streptokinase. This fusion
 CC protein was used in the design of modified streptokinase has an in vitro
 CC degradation rate at least 2 times slower than that of native
 CC streptokinase. Compounds containing modified streptokinases are
 CC specifically used as thrombolytic agents for dissolving blood clots in
 CC vivo in a mammal, preferably at a dose of 2000 U/Kg, optionally as a
 CC bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 1194 AA;
 Alignment Scores:
 Pred. No.: 7,46e-192 Length: 1194
 Score: 2150.50 Matches: 420
 Percent Similarity: 97.68% Conservative: 1
 Best Local Similarity: 97.45% Mismatches: 5
 Query Match: 68.62% Indels: 5
 DB: 2 Gaps: 3
 US-09-940-235-11 (1-1782) x AAW21726 (1-1194)
 QY 496 CAGACCAATCGAC-----GATCTGGCCCTTACCGATGTTGTAATGCTGGA 546
 Db 366 GlnThrAsnSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
 QY 547 CCTGAGTGGCTGCTTACACCGTCCATCTGTCAACAACCAATGTTGTTAGCTTCT 606
 Db 385 ProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValIleSerValAla 404
 QY 607 GGTACTGTGAGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATGATCTACA 666
 Db 405 GlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIleAspLeuThr 424
 QY 667 TCAAGACCTGCTCAAGAGAAAGACAGAGCAAGCTTAACTCCAAATCAAAATCAATT 726
 Db 425 SerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSerTyrPhe 444
 QY 727 GCTACTGATAGTGGCGGATGTCACATTAACCTTGAAGAAGCTGACTTAAGGCTATT 786
 Db 445 AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLysAlaIle 464
 QY 787 CAAGAATAATGATCGGTAACGTCACAGTAACAGACACTTGTGAGTCATGATTTT 846
 Db 465 GlnGluGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluValIleAspPhe 484
 QY 847 GCAAGCATGCAACCATTAATGATCGAAGAGCGCAAGGCTTACTTGTGACCAAGATGAT 906
 Db 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAspValAspGly 504
 QY 907 TCGGTAACTTGGCGAACCCAACTGTCCAAAGATTTTGTAGCGGACATGTGGCGGT 966
 Db 505 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 524
 QY 967 AGACCAATTAAGAAAAACAATCAACAAACCAAGCGAAATCTGTGATGTGGAATTAATCT 1026
 Db 525 Arg---TyrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 543
 QY 1027 GTACAGTTTACTCCCTTAAACCTGATGACGATTTGACACAGGCTCAAGATATCAAG 1086
 Db 544 ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLys 563

```

QY 1087 CTATTGAAAACCTAGCTATCGGTGACACCATCATCTCAAGAACTTACTAGCTCAAGCA 1146
Db 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuAlaGlnAla 583
QY 1147 CAAGCATTGTTAAACAAAACCAACCGAGCTATACGATTATGAAAGTGAATCTCTCAATC 1206
Db 584 GlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrClnLysAspSerSerIle 603
QY 1207 GTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATGATCAAGAGTTACTTAC 1266
Db 604 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetCaspGlnGluPheThrTyr 623
QY 1267 CGTGTGTTAAATATCGGGAACACGCTTATAGATCATATAAATAATCTGCTGTAATGAAGA 1326
Db 624 ArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysSerClnLysLeuGlnGlu 643
QY 1327 ATTAACACACCTGACCTGATCTCTGAGAAATATTACGCTCTTAAAGGCGAAAGCCG 1386
Db 644 IleAsnAsnThrAspLeuLysSerGlnLysTyrTyrValLeuLysGlnLysPro 663
QY 1387 TATGATCCCTTGTGATCGACGATCACTTGAACGTTGACCATGATGATGTCGAT 1446
Db 664 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 683
QY 1447 ACCAAGCAATGCTGCTAAAGAGTGAAGAGCTCTTACAGCTAGCGAAGCTTACGCTTC 1506
Db 684 ThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerClnLysAsnLeuAspPhe 703
QY 1507 AGAGATTATACGATCTCTGTATGAAGCTTAAAGCTTACCTTACCAACATCTCGATGCTTTT 1566
Db 704 ArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeuAspAlaPhe 723
QY 1567 GGTATTATGACTATATACCTTAACTGGAAGAGTGAAGAGTGAAGTGAAGCAACCGT 1626
Db 724 GlyIleMetAspPyrThrLeuThrGlyLysValGlnAspAsnHisAspThrAsnArg 743
QY 1627 ATCTAATACCGTTTATATGCGCAAGAGCAACCGAAGAGAGAAATGCTATCATTTAGCC 1686
Db 744 IleIleThrValTyrMetClnLysArgProGlnGlnLysAsnHisAspTyrHisLeuAla 763
QY 1687 TATGATTAAGATGCTTATACCGAAGAGAGAAAGAGTTTACGCTACCTGCTTATACA 1746
Db 764 TyrAspLysAspArgTyrThrGlnGlnGlnValArgGlnValTyrSerTyrLeuArgTyrThr 783
QY 1747 GGAGACCTATATACCTGATTAACCTTACGCAAA 1779
Db 784 GlyThrProIleProAspAsnProAsnAspLys 794

RESULT 13
AAR12891
ID AAR12891 standard; protein; 499 AA.
XX
XX AAR12891;
AC
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-SEP-1991 (first entry)
DE Streptokinase fused to yeast alpha factor secretion sequence.
XX
XX Fusion protein; blood clotting; coagulation; fibrinolysis;
XX antithrombotic; chromolysis; streptokinase.
XX
OS Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 100009.
XX
XX Key Location/Qualifiers
XX Peptide 1..85
XX FT /label= pre-pro alpha factor secretion sequence
XX FT /note= "S. cerevisiae"
XX FT Protein 86..499
XX /label= mature streptokinase
XX
XX MO9109125-A.

```

```

XX
XX 27-JUN-1991.
PD
XX
XX 07-DEC-1989; 89GB-00027722.
PF
XX
XX 07-DEC-1989; 89GB-00027722.
PR
XX 07-DEC-1990; 90MO-GB001911.
XX
XX (BRB1-) BRITISH BIO-TECHNOLOGY LTD.
XX
XX Dawson KM, Hunter MG, Czapliewsk LG;
PI
XX
XX WPI; 1991-208151/28.
DR
XX N-PSDB; AAQ12158.
XX
XX Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT fractions having greater antithrombotic activity for therapy and
PT prophylaxis.
XX
XX Disclosure; Page 86; 115pp; English.
PS
XX
XX The streptokinase sequence was obtd. from PCR amplified chromosomal DNA
CC from S. equisimilis (Lancefield's Gp C) ATCC 10009 or ATCC 9642 (the
CC primers used for the PCR were based on the published DNA sequence of S.
CC equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.D., Gene 34
CC 357-362 (1985)). The gene was fused to DNA encoding the yeast alpha
CC factor pre-pro-secretion sequence in an expression vector, pSMD1/152, for
CC prodn. of recombinant strepto- kinase in S. cerevisiae strain BJ2168. See
CC also AAR12887-R12889, AAR12892-R12894, AAR12885 and AAR12522. (Updated on
CC 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise
CC OS field)
XX
XX
SQ Sequence 499 AA;
XX
XX
Alignment Scores:
Pred. No.: 2,62e-191 Length: 499
Score: 2143.00 Matches: 413
Percent Similarity: 96.29% Conservative: 2
Best Local Similarity: 95.82% Mismatches: 16
Query Match: 68.38% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x AAR12891 (1-499)
QY 487 ACCTGTGTGACGACCATGATCGGCCCCCTTACCGATGTTGTTATGCTGCA 546
Db 69 ThrIleAlaSerIleAlaIleAlaLysGlnGlnGlnValSerLeuAspLysArgIleAlaGly 88
QY 547 CCTGAGTGGCTGTGTAACCGTGCATCTGTCAACAGCAAGCAATGTTGTTAGCGTGTCT 606
Db 89 ProGlnThrPheLeuAsnAspArgProSerValAsnAspSerGlnLeuValValSerValAla 108
QY 607 GGTACTGTTGAGGGGAGCAATCAAGCATTAGTCTTAAATTTTGGAAATCGATCTTACA 666
Db 109 GlyThrValGlnGlnLysThrAsnGlnAspIleSerLeuLysPheGlnIleAspLeuThr 128
QY 667 TCAGCACTGCTGTATGAGGAGAAAGACAGCAGAGCTTAAAGTCCAAATCAAAACCTTT 726
Db 129 SerArgProAlaHisGlyGlyLysTrpGlnGlnGlnLysSerProLysSerLysProPhe 148
QY 727 GGTACTGATATGAGGGGCGATGTGCACTTAACTGAGAAAGCTGACTTAAAGGCTATT 786
Db 149 AlaThrAspSerIleValMetProHisLysLeuGlnLysAlaAspLeuLysAlaIle 168
QY 787 CAAGAACATTTGATCGCTAAGCTTCAACAGTACAGACGACTTATGAGTCAATGATTTT 846
Db 169 GlnGlnGlnLeuLeuIleAsnValHisSerAsnAspAspTyrThrGlnValIleAspPhe 188
QY 847 GCAAGCGATGCAACCATTAAGTCAAGAAAGCGAAGGCTTACTTGTGCAAAAGATGT 906
Db 189 AlaSerAspAlaThrIleThrAspArgAsnGlnLysValTyrPheAlaAspLysAspGly 208
QY 907 TCGGTAACTTGCCGACCAACCTGTCCAAAGATTTTGTAAAGCGGACATGTCCGCGTT 966

```

```

Db      209  SerValThLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 228
Qy      967  AGACCATATAAGAAAAACCAATATACAACCAAGCGAAATCTGTGATGGAATATATCT 1026
Db      229  ArgProTylLysGluSerProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 248
Qy      1027  GTAAGATTACTCCCTTAAACCTGATGACGATTTCCAGACCAAGCTCTCAAGATACTAG 1086
Db      249  ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLys 268
Qy      1087  CTAATTGAAAACTAGTACTCGGTGACACATGACATCTCAAGAAATTACTAGCTCAGCA 1146
Db      269  LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuAlaGlnAla 288
Qy      1147  CAAGCATTTTAAACAAAAACCAAGCGCTATACGATTTATGAAAGCTGACCTCAATC 1206
Db      289  GlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluLysAspSerIle 308
Qy      1207  GTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGATTACTTAC 1266
Db      309  ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyr 328
Qy      1267  CGTGTTAAAAATCGGGAACAGCTTATAGATCAATAAAAATCGCTGTGATGAAAGAA 1326
Db      329  HisValLysAsnArgGluGlnAlaTyrGluIleAsnLysSerGlyLeuAsnGluGlu 348
Qy      1327  ATAAACAACAGTACGCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGGAAGAACG 1386
Db      349  IleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGlyGluLysPro 368
Qy      1387  TATGATCCCTTTGATGCGAGTCACTTGAAACTGTCACCATCAATCTGTGATGTGCAT 1446
Db      369  TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsn 388
Qy      1447  ACCAAGCAATTGCTAAAAAGTAGAGCAAGCTCTTAAACAGTACGAAAGCTTAAAGCTTC 1506
Db      389  ThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsnLeuAspPhe 408
Qy      1507  AGAGATTTTATAGATCTCGTGATAAGGCTTAAAGCTCTACCAACATCTCGATGCTTTT 1566
Db      409  ArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysAsnAsnLeuAspAlaPhe 428
Qy      1567  GGTATTTATGACATATCTTAACTGAAAAAGTAGAGATATACAGATGACCAACCGT 1626
Db      429  GlyIleMetAspTyrThrIleuThrGlyLysValGlnAspAsnHisAspAspThrAsnAsp 448
Qy      1627  ATCATTAACCGTTATATGGGCAAGCGACCCGAAAGAGAGATGCTAGTACATTTAGCC 1686
Db      449  IleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyrHisLeuAla 468
Qy      1687  TATGATTAAGATCGTTATATCCGAAAGAAAGAGAAATTACAGCTACCTGCGTTATACA 1746
Db      469  TyrAspLysAspArgTyrThrGluGluGluArgLysGluValTyrSerTyrLeuAspTyrThr 488
Qy      1747  GGGACACCTATACCTGATTAACCTTAAGCAAAA 1779
Db      489  GlyThrProIleProAspAsnProAsnAspLys 499

RESULT 14
ID      AAW21728
AAW21728 standard; protein; 813 AA.
XX
AC      AAW21728;
XX
DT      17-OCT-2003 (revised)
DT      01-OCT-1997 (first entry)
XX
DE      Wild type plasminogen-binding fragment of Streptokinase.
XX      Plasminogen-binding fragment; streptokinase; degradation; MBP;
KW      chromolytic agent; blood clot; bolus; maltose-binding protein.
XX

```

```

OS      Streptococcus dysgalactiae subsp. equisimilis.
XX
PN      WO9641883-A1.
XX
PD      27-DEC-1996.
XX
PF      07-JUN-1996; 96WO-US009640.
XX
PR      09-JUN-1995; 95US-00488940.
XX
PA      (HARD ) HARVARD COLLEGE.
XX
PI      Reed GL;
XX
DR      WPI; 1997-065469/06.
XX
PT      Modified forms of streptokinase resistant to enzymatic cleavage - useful
PT      as chromolytic agents in treating thrombosis and in medical equipment.
XX
PS      Example 1; Page 12-13; 65pp; English.
XX
CC      This sequence represents the wild type plasminogen-binding fragment of
CC      streptokinase. This fragment was used in the design of a modified
CC      streptokinase has an in vitro degradation rate at least 2 times slower
CC      than that of native streptokinase. Compounds containing modified
CC      streptokinases are specifically used as chromolytic agents for
CC      dissolving blood clots in vivo in a mammal, preferably at a dose of 20000
CC      U/kg, opt. as a bolus rather than by continuous infusion. (updated on 17-
CC      OCT-2003 to standardise OS field)
XX
SQ      Sequence 813 AA;

Alignment Scores:
Pred. No.: 5.5e-191 Length: 813
Score: 2140.50 Matches: 413
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 0
Query Match: 68.30% Indels: 1
DB: 2 Gaps: 1

US-09-940-235-11 (1-1782) x AAW21728 (1-813)
Qy      538  ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACCAACAGCAATGGTGT 597
Db      1  IleAlaGlyProGluThrProLeuLeuAspArgProSerValAsnSerGlnLeuValAl 20
Qy      598  AGCGTTGCTGTACTGTGAGGGGACCAATCAAGACATTAATTTTGAATTC 657
Db      21  SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
Qy      658  GATTAACATCAGCACTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 717
Db      41  AspLeuThrSerArgProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 60
Qy      718  AAACCATTTGCTACTGATAGTGGCGAGATGTCATTAACCTTGAAGAAAGCTTACTA 777
Db      61  LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
Qy      778  AAGCTATTCAAGACATTAATGATCGTAACTGACAGTAAAGCAAGTACTTGTGAGTGC 837
Db      81  LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
Qy      838  ATTGATTTTSCAAGCGATGCAACCATTAATGATGAAACGGCAAGGTCTACTTTGTGAC 897
Db      101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Qy      898  AAAGATGTTTCGGTAACTTGCCGACCAACCTGTCCAAAGATTTTGTAAAGCGACAT 957
Db      121  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
Qy      958  GTGCGGCTTGAACCATTAAGAAACCAATACAAACCAAGGAATCGTTGATGAGTG 1017
Db      141  ValArgValArg--TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159

```

QY 1018 GAATATACGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGGTCAAA 1077
 DB 160 GtutyrtrthvalGlnPheThrProLeuAenProAspAspAspPheArgProGlyLeuLys 179
 QY 1078 GATCTAAGCTATTTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 1137
 DB 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleLeu 199
 QY 1138 GCTCAAGCAAGACATTTTAAACAAACCCAGCGCTATACGATTATGAAAGCTGAC 1197
 DB 200 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyThrThrIleTyrgluAsp 219
 QY 1198 TCCTCAATCGTCACTCATGACATGATCTTCCGTACGATTTTACCAATGATCAAGAG 1257
 DB 220 SerSerIleValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnIle 239
 QY 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAATAATCGTCTG 1317
 DB 240 PheThrTyrgValValLysAsnArgGlnIleAlaTyArgIleAsnLysLysSerGlyLeu 259
 QY 1318 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1377
 DB 260 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyTyValLeuLysLysGly 279
 QY 1378 GAAAAGCGGTATGATCCCTTGTATGCGACATCTTGAACCTGTCACCAATCAATACGTT 1437
 DB 280 GlnLysProTyraAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 299
 QY 1438 GATGTGATACCAACGAATTTGCTAAATAAGTGACAGCTCTTAAACAGTACGAACCTAAC 1497
 DB 300 AspValAspThrAsnGlnLeuLeuLysSerGlnIleLeuThrIleAspGlnIleAsn 319
 QY 1498 TTTAGCTTCAAGATTTATACATCTCTGATTAAGGCTTAACTCTTCAACAATCTTC 1557
 DB 320 LeuAspPheArgAspLeuTyraAspProArgAspLysAlaLysLeuLeuTyraAsnLeu 339
 QY 1558 GATGCTTTTGGTATATGACATCTTACCTTAACTGGAATAAGATATATCCAGTAC 1617
 DB 340 AspAlaPheGlyLysLeuAspTyThrLeuThrGlnLysValGlnAspAsnHisAspAsp 359
 QY 1618 ACCAACCGTATCATACCGTTATATGCGCAAGCGAACGGAAGAAATGCTTACCTAT 1677
 DB 360 ThrAsnArgIleIleThrValTyMetGlyLysArgProGlnIleGlnAsnAlaSerTy 379
 QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAAAGAGAGCTTTACAGCTTAC 1737
 DB 380 HisLeuAlaTyraAspLysAspArgTyThrGlnGlnIleValTySerTyLeu 399
 QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAA 1779
 DB 400 ArgTyThrGlnTyThrProIleProAspAsnProAsnAspLys 413
 RESULT 15
 AAM86143
 ID AAM86143 standard, protein; 414 AA.
 AC AAM86143;
 DT 17-OCT-2003 (revised)
 DT 03-MAR-1999 (first entry)
 XX Streptokinase (SK) protein sequence.
 DE Streptokinase (SK) protein sequence.
 KW Non-immunogenic; epitope; T-cell; immunogenicity; immune system; SK;
 KW immunogly bulin; therapeutic; streptokinase.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 XX MO9852976-A1.
 PN 26-NOV-1998.
 PD
 XX

PF 21-MAY-1998; 98MO-GB001473.
 XX
 PR 21-MAY-1997; 97GB-00010480.
 PR 31-JUL-1997; 97GB-00016197.
 PR 28-NOV-1997; 97GB-00025270.
 PR 02-DEC-1997; 97US-0067235P.
 PR 14-APR-1998; 98GB-00007751.
 XX
 PA (BIOV-) BIOVATION LTD.
 XX
 PI Carr FU;
 XX
 DR WPI; 1999-045301/04.
 XX
 PT Reducing immunogenicity of proteins - by modifying the amino acid
 PT sequence of the protein to eliminate potential epitopes for T-cells of a
 PT given species.
 PS
 XX Example 6; Fig 28; 77pp; English.
 CC The invention relates to a method for the production of non-immunogenic
 CC proteins. The method comprises determining at least part of the amino
 CC acid sequence of the protein; (b) identifying in the amino acid sequence
 CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
 CC species; and (c) modifying the amino acid sequence to eliminate at least
 CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
 CC reduce the immunogenicity of the protein when exposed to the immune
 CC system of the given species. A method of analysing a pre-existing protein
 CC to predict the basis for immunogenic responses is also provided. The
 CC methods can be used particularly for reducing the immunogenicity of
 CC immunoglobulins or therapeutic proteins, e.g. Streptokinase (SK). The
 CC products can be used for diagnosis and therapy. The present sequence
 CC represents the amino acid sequence of the SK protein. (updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 2,6e-190 Length: 414
 Score: 2132.00 Matches: 409
 Percent Similarity: 99.03% Conservative: 1
 Best Local Similarity: 98.79% Mismatches: 4
 Query Match: 68.03% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-11 (1-1782) x AAM86143 (1-414)
 QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTGTTGTT 597
 DB 1 IleAlaGlyProGlnIleProLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 598 AGCGTGTCTGTACTGTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGAATC 657
 DB 21 SerValAlaGlyThrValGlnGlnTyThrAsnGlnAspIleSerLeuLysPhePheGlnIle 40
 QY 658 GATCTAATCATGACGACCTGCTCATGAGAGAAAGACAGACGAAAGCTTAAAGTCCAAATCA 717
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnIleLeuSerProLysSer 60
 QY 718 AAACATTTGCTACTGTATGATGGGCGGATGTCACATTAACCTTGGAAGCGACTTACTA 777
 DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 80
 QY 778 AAGGCTATTCAAGAACATTTGATCGCTTACGTCACAGTCAAGACGACTTGTGAGGTC 837
 DB 81 LysAlaIleGlnGlnIleLeuIleAlaAsnValHisSerAsnAspArgTyPheGlnVal 100
 QY 838 ATTGATTTTCAAGACGATGCAACCATTAATCTGATCGAAAGCGCAAGGTCTACTTGTGAC 897
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlnLysValTyPheAlaAsp 120
 QY 898 AAAGATGCTTGGTAACTTGGCGACCACTGCTCAAGAAATTTTGTAAAGGGAAT 957

Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
QY 958 GTGCGGTTAGACCATATAAAGAAAAACAATACAAACCAAGCAAAATCTGTGATGTC 1017
Db 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATACGTGTACAGTTTACTCCCTTAACCCCTGATGACGATTTGAGACCAAGTCTCAA 1077
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATCTAAGCTATTGAAACACATAGCTATGGTGACACCATCATCTCAAGATTACTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY 1138 GCTCAGACCAAGCATTTTAAACAAACCAACCAAGCTATACGATTTTATGAAAGTGAC 1197
Db 201 AlaglnAlaglnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 220
QY 1198 TCCTCAATGTCATCATGACAAATGACATTTTCGGTACGATTTTACCAATGATCAAGAG 1257
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1258 TTTACTTACCGGTATTAATAATCGGAAACAGCTTATAGATCAATATAAAATCTGTCGTG 1317
Db 241 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnLysSerGlyLeu 260
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTCGAATAATTAACGTCTTAATAAAAGGG 1377
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 280
QY 1378 GAAAACCGTATGATCCCTTGTATGTCAGTCACTTGAACGTTCACCATCAATACGTT 1437
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1438 GATGTCGATACCAACGAATGCTAATAAGTAGACAGCTTTAACAGCTACGAAAGTAAAC 1497
Db 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 320
QY 1498 TTAGACTTCAGAGATTATACGATCTCCGATGAAGGCTAACTACTACACAAATCTC 1557
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1558 GATGCTTTGGTATTATGACACTATACCTTAACGTGAAAAGTAGAGATAATCAAGATGAC 1617
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1618 ACCAACCGTATCATAAACGTTTATATGAGGCAAGCAACCCGAGAGAGAAATGCTAGCTAT 1677
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAsnAlaSerTyr 380
QY 1678 CATTTAGCCTATGATTAAGAATGCTTATACCGAAGAGAAAGAGAGATTACAGTACCTG 1737
Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnLysValTyrSerTyrLeu 400
QY 1738 CGTTATACGAGGAGCACTATACCTGATACCCCTTAACGACAA 1779
Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414

Search completed: January 28, 2006, 02:08:39
Job time : 145.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 25.4212 Seconds
(without alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134

Sequence: 1 tcgcctcaagtcgctcgcsy.....ataccctaagacaataa 1782

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=rlp
-O=/cgn2_1/USFTO_spool_p/US09940235/runat_27012006_144219_27602/app_query.fasta_1.7708
-DB=Issued Patents_AA -QFMT=fastan -SUPFIX=ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=100 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1_1.159 @runat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5/COMB.pep:*
2: /cgn2_6/prodata/1/iaa/6/COMB.pep:*
3: /cgn2_6/prodata/1/iaa/7/COMB.pep:*
4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep:*
5: /cgn2_6/prodata/1/iaa/RA.COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2168	69.2	795	2	US-09-211-542A-2
2	2165	69.1	483	1	US-07-854-596B-43
3	2158	68.9	414	2	US-09-211-542A-6
4	2158	68.9	414	6	5240845-1
5	2158	68.9	440	1	US-08-488-940-1
6	2153	68.7	413	2	US-10-360-101-264
7	2150.5	68.6	1194	1	US-08-488-940-1
8	2143	68.4	499	1	US-07-854-596B-28
9	2140.5	68.3	813	1	US-08-488-940-3
10	2132	68.0	415	1	US-07-854-596B-26
11	2132	68.0	435	1	US-07-854-596B-19
12	2132	68.0	440	1	US-07-854-596B-15

13	2132	68.0	483	1	US-07-854-596B-47	Sequence 47, Appl
14	2132	68.0	859	1	US-07-854-596B-35	Sequence 35, Appl
15	2120.5	67.7	1194	1	US-08-488-940-17	Sequence 17, Appl
16	2114.5	67.5	1194	1	US-08-488-940-18	Sequence 18, Appl
17	2090.5	66.7	413	1	US-08-759-599-12	Sequence 12, Appl
18	2090.5	66.7	413	1	US-09-294-457-12	Sequence 12, Appl
19	2090.5	66.7	413	2	US-09-919-703-12	Sequence 12, Appl
20	2071	66.1	800	1	US-08-488-940-4	Sequence 4, Appl
21	2071	66.1	1181	1	US-08-488-940-2	Sequence 2, Appl
22	2061	65.8	401	2	US-09-374-038-1	Sequence 1, Appl
23	2061	65.8	401	2	US-09-658-179-1	Sequence 1, Appl
24	2061	65.8	413	2	US-09-374-038-2	Sequence 2, Appl
25	2061	65.8	413	2	US-09-658-179-2	Sequence 2, Appl
26	1902	60.7	384	2	US-09-374-038-4	Sequence 4, Appl
27	1902	60.7	384	2	US-09-658-179-4	Sequence 4, Appl
28	1898	60.6	372	2	US-09-374-038-3	Sequence 3, Appl
29	1898	60.6	372	2	US-09-658-179-3	Sequence 3, Appl
30	1876	59.9	747	1	US-07-854-596B-40	Sequence 3, Appl
31	1875	59.8	369	1	US-07-854-596B-31	Sequence 40, Appl
32	1872.5	59.7	736	2	US-09-211-542A-4	Sequence 31, Appl
33	1860	59.3	356	2	US-09-211-542A-12	Sequence 12, Appl
34	785	25.0	150	2	US-09-211-542A-14	Sequence 14, Appl
35	504.5	16.1	2324	1	US-08-283-857-1	Sequence 1, Appl
36	504.5	16.1	2324	4	PCT-US95-09819-1	Sequence 1, Appl
37	504.5	16.1	2446	1	US-08-551-356-2	Sequence 2, Appl
38	504.5	16.1	2446	4	PCT-US93-12687-2	Sequence 2, Appl
39	503	16.0	2231	1	US-08-153-799-16	Sequence 16, Appl
40	503	16.0	2335	2	US-10-360-101-235	Sequence 235, App
41	503	16.0	2386	1	US-09-016-366A-12	Sequence 12, Appl
42	503	16.0	2386	2	US-09-961-403-1	Sequence 1, Appl
43	499.5	15.9	2327	6	5455158-1	Parent No. 5455158
44	283	9.0	286	2	US-09-078-091-4	Sequence 4, Appl
45	275	8.8	286	2	US-09-078-091-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:

; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP

; STREET: 125 Summer Street
; CITY: Boston

; STATE: Massachusetts
; COUNTRY: USA

; ZIP: 02110

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION NUMBER: US/09/211.542A
; FILING DATE: 15-December-1998

; CLASSIFICATION: 1653

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497

; FILING DATE: 15-December-1997

; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.

; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 443-9292

; TELEFAX: (617) 443-0004
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 795 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-211-542A-2

Alignment Scores:

Alignment No.:	5, 2e-211	Length:	795
Score:	2168.00	Matches:	421
Percent Similarity:	97.918	Conservative:	1
Best Local Similarity:	97.688	Mismatches:	5
Query Match:	69.184	Indels:	2
		Gaps:	2

US-09-940-235-11 (1-1782) x US-09-211-542A-2 (1-795)

```

QY 496 CAGACCAATCGAGC-----GGATCTGGCCCTTCACCGAGTTCGTATTCGTGA 546
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 GlnThrAsnSerSerValProGlyArgGlySerIle--GlnGlyArgIleAlaGly 384

QY 547 CCTGAGTGGCTCTAGACCGCTCATCTGTTCACACAAAGCCAAATTGGTTGACCGTTGCT 606
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 385 ProGlnTrpLeuAsnAspArgProSerValAsnAsnSerGlnLeuValValSerValAla 404

QY 607 GGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATAATCGATCAACA 666
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 405 GlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGlnIleAspLeuThr 424

QY 667 TCACGACCTGCTCATGAGAGAAAGACAGACGAAAGCTTAAGTCCAAATCAAAACGATT 726
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 425 SerArgProAlaHisGlyGlySerThrGlnGlnGlyLeuSerProLysSerLysProPhe 444

QY 727 GCTACTGATAGTGGCCGACATGTCACATTAACCTTGAGAAAGCTGACTTAAGGCTATT 786
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 445 AlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLysAlaIle 464

QY 787 CAAGAACAATGATGCTTAACGTCACAGTAAACGACGACTTGAAGTCATTAATTTT 846
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 465 GlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspLysPheGlnValIleAspPhe 484

QY 847 GCAGCGATGCAACCATTAAGTCAAGAAACGCAAGGTCCTACTTGTCTGACAAAGATGAT 906
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValLysPheAlaAspLysAspGly 504

QY 907 TCGGTAACCTTGCCGACCAACCTGTCACAAGAATTTTGTCTTAAGCGCAATGTCGCCGT 966
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 505 SerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHisValArgVal 524

QY 967 AGACCATTAAGAAAAACCAATACAAACCAAGCAAGTCTGTTGATGCAATTAATTAAT 1026
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 525 ArgProLysLysGlnLysProIleGlnAsnGlnAlaLysSerValAspValGlnLysThr 544

QY 1027 GTACAGTTTACCTCCCTTAACCTGATGACATTTTGAAGACAGGTCTCAAGATTAAG 1086
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 545 ValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLysAspThrLys 564

QY 1087 CTATTGAAAAACCTAGCTATGCGTGAACACATCAATCTCAAGATTAATTAAGTCAAGCA 1146
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 565 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuAlaGlnAla 584

QY 1147 CAAGCATTTTAAACAAAACCAACCGGCTATACATTTTGAAGTGAATCTCTCAATC 1206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 585 GlnSerIleLeuAsnLysAsnHisProGlyLysThrIleLysGlnLysAspSerSerIle 604

QY 1207 GTCACTCATGACATGACATTTTCCGTACGATTTTCAATGATGACAAAGTTTACTTAC 1266
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 605 ValThrHisAspAsnAspIlePheAspGlnThrIleLeuProMetAspGlnLysPheThrLys 624

QY 1267 CGTGTTAAAAATCGGGAACAAAGCTTATAGATCAATTAATAAATCTGCTGTAAGAAGA 1326
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 625 ArgValLysAsnArgGlnGlnAlaLysArgIleAsnLysLysSerGlyLeuAsnGlnGln 644

QY 1327 ATAAACAACACATGACCTGATCTCTAGAAATATTAGCTCTTAATAAAGGGAAGCCG 1386

```

```

Db 645 IleAsnAsnThrAspLeuIleSerGlnLysLysThrValLeuLysLysGlnLysPhe 664

QY 1387 TATGATTCCTTTGATCGCACTGACCTGAAACCTGTCACCATCAATACGTTGATGAT 1446
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 665 TyrAspProPheAspArgSerHisLysLeuLysPheThrIleLysThrValAspValAsp 684

QY 1447 ACCAAGCAATTTGCTAAATAAAGAGACGACGCTTAAACACTGACGCAAGCTTAAGCTTC 1506
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 685 ThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsnLeuAspPhe 704

QY 1507 AGAGATTTATACGATCTCGTGAATAGGCTTAACCTGACCTACCAACAACTCGATGCTTT 1566
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 705 ArgAspLeuLysThrAspProAlaGlyAspLysAlaLysLeuLeuLysThrAsnLeuAspAlaPhe 724

QY 1567 GGTAATTATGACATATACCTTAATCTGAGAAAGTGAAGATATACGATGACCAACCGT 1626
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 725 GlyIleThrValThrMetGlyLysArgProGlnGlyGlnLysAlaSerThrHisLeuAla 744

QY 1627 ATGATAACGCTTATATAGGCAAGCCGACCGAAGAGAGAAATGCTACTATCATTTAGCC 1686
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 745 IleIleThrValThrMetGlyLysArgProGlnGlyGlnLysAlaSerThrHisLeuAla 764

QY 1687 TATGATTAAGATGCTTATACCGAAGAAAGACGAAAGTTTACAGTCTGCGTTATACA 1746
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 765 TyrAspLysAspArgLysThrGlnGlnGlnLysGlnValLysThrLysLeuArgLysThr 784

QY 1747 GGGACACTTATACCTGATTAACCTTAACGACAAA 1779
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 785 GlyThrProIleProAspAsnProAsnAspLys 795

RESULT 2
US-07-854-596B-43
; Sequence 43, Application US/07854596B
; Patent No. 5434073
;
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SBO ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-43

```

Alignment Scores:

Pred. No.: 8,15e-211 Length: 483
 Score: 2165.00 Matches: 434
 Percent Similarity: 85.22% Conservative: 10
 Best Local Similarity: 83.30% Mismatches: 27
 Query Match: 69.08% Indels: 50
 DB: 1 Gaps: 9

US-09-940-235-11 (1-1782) x US-07-854-596B-43 (1-483)

244 ACAGATTGACCATGAGTGGAGAGTGTTCATGCTGGGACTTCCTATGNGT 303
 4 ThrAspCysThrGluSerGlyGln-----AsnLeuCys-----LeuCys--- 16
 304 CGAGAAACGTGGAGAGGACGACGACGATCATCTTGACATTCTAGAATATGATGAC 363
 17 -----GluGlySerAsn-----ValCysGlyGlnGlyAsnLysCys--- 28
 364 GATCAGACACAAGACATCTTATAGAAATGAGACACCTGGAGCAAGAAAGATTAATCGA 423
 29 -----IleLeuGlySerAspGlyLysLysAsn----- 37
 424 GGAACCTGCTCCAGTCATCTGCACAGGACGCGGAGAGAGTGAAGTGAAGC 483
 38 -----GlnCysVal-----ThrGlyGluGly----- 44
 484 CACACCTCTGTGACAGCACATCGACGCGATCTGCGCCCTTCACCGATGTT----- 534
 45 -----ThrProLysProGlnSerHisAsnAspGlyAspPheGluGlnIleProGluGln 62
 535 -----CGTATTGCTGACCTGAGTGGCTGCTGACGCTTCATCTGTC 576
 63 TyrLeuGlnIleGluGlyArgIleAlaGlyProGluIleProLeuLeuAspArgProSerVal 82
 577 AACACAGCGCAATGGTGTGACCGTTCGCTGACTGCTTGAAGGAGCAATCAAGACATT 636
 83 AsnAsnSerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIle 102
 637 AGTCTTAATTTTGAATTCGATCAATCAATCAAGACCTGCTCATGGAGAGAAAGACAG 696
 103 SerLeuLysPhePheGlnIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGln 122
 697 CAAGGCTTAAGTCAAAATCAAAACCAATTTGCTACTGATAGTGGCGGATGCATAA 756
 123 GlnGlyLeuSerProLysSerLysProPheAlaThrAspSerIleAlaMetProHisLys 142
 757 CTGGAAGAAAGCTGACTTACTTAAGGCTATTCAGAAACAATGATCGCTAACGTCACAGT 816
 143 LeuGluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSer 162
 817 AACGACGACTACTTGAAGTCAATGATTTTGGACAGGATGCACCACTTACTGATCGAAC 876
 163 AsnAspAspTyrPheGlnValIleAspPheAlaSerAspAlaThrIleThrAspArgAsn 182
 877 GCGAAGTCTACTTGTGTCGACAAAGATGGTTCGTTACCTGGCGGACCCGAACTGTCCAA 936
 183 GlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGln 202
 937 GAATTTTGTCTAAGCGACATGTGCGCTTGAACCATATTAAGAAACCAATATCAAAAC 996
 203 GluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsn 222
 997 CAAGCAAAATCTGTGATGTGGAATATCTGATCACTTACTCCCTTAAACCTGATGAC 1056
 223 GlnAlaLysSerValAspValIleTyrThrValGlnPheThrProLeuAsnProAspAsp 242
 1057 GATTTCAGACGAGGTCTCAAGATCTAAGCTATTGAAAAACCTACTATCGGTGACACC 1116
 243 AspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThr 262
 1117 ATCAATCTCAAGAAATTAATCTACCTCAAGCAAAAGATTTTAAACAAACCAACGAGC 1176
 263 IleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGly 282

1177 TATACGATTTTGAACGTAAGTCTCTCAATCGTCACTCATGACAAATGATTTTCGTAGC 1236
 283 TyrThrIleTyrGluAlaGAspSerSerIleValIleThrHisAspAsnAspIlePheArgThr 302
 1237 ATTTTACCAATGATCAAGACTTTTACTTACCGGTGTTAAATTCGGAACAAAGCTTATAGC 1296
 303 IleLeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluAlaIleTyrGlu 322
 1297 ATCAATAAAAAATCGTCTGAATGAAGAAATTAACAACATGACCTGATCTTCGAGAA 1356
 323 IleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLys 342
 1357 TATTACGTCCTTAAAAAAAGGGAAGCGGATATCCCTTGTGATCGAGTCACTTGA 1416
 343 TyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLys 362
 1417 CTGTTCAACATCAAAATAGTTGATGTGATACCAACGAATTCCTAAAGTGAAGCAGCTC 1476
 363 LeuPheThrIleLysTyrValAspValAsnThrAsnGlnLeuLeuLysSerGluGlnLeu 382
 1477 TTAACTGATGCGCAAGCTTACTTGAATTCAGAGATTTATACGATCTCTGATTAAGCT 1536
 383 LeuThrAlaSerGluAlaGAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAla 402
 1537 AAATCTACTTACCAACAATTCCTGATGCTTTGGTATTATGACTATTAATCTTAACCTGGA 1596
 403 LysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLys 422
 1597 GTAGAGATATATACGATGACACCAACCGTATCAATAACGTTTATATGAGCAAGCAGCC 1656
 423 ValGluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgPro 442
 1657 GAAGGAGAAATGCTACTATCATTTAGCCTATGATTAAGATGCTTATACGGAAGAA 1716
 443 GluGlyLysAsnAlaSerTyrHisLeuValAlaTyrAspLysAspArgTyrThrGluGluGlu 462
 1717 CGAAGATTTTACGCTACCTGCGCTTATACAGGAGCACCTATACCTGATTAACCTTAAGC 1776
 463 ArgGluValTyrSerTyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAsp 482
 1777 AAA 1779
 483 Lys 483

RESULT 3
 US-09-211-542A-6
 : Sequence 6, Application US/09211542A
 : Patent No. 6210667
 : GENERAL INFORMATION:
 : APPLICANT: Reed, Guy L.
 : TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BROMBERG & SUNSTEIN, LLP
 : STREET: 125 Summer Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: USA
 : ZIP: 02110
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/211,542A
 : FILING DATE: 15-December-1998
 : CLASSIFICATION: 1653
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/069,497
 : FILING DATE: 15-December-1997
 : ATTORNEY/AGENT INFORMATION:
 :

NAME: Attorney, Strimpel, Harriet M.
 REGISTRATION NUMBER: 37,008
 REFERENCE/DOCKET NUMBER: 1874/111
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)443-9292
 TELEFAX: (617)443-0004
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 414 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-211-542A-6

Alignment Scores:

Pred. No.:	3,87e-210	Length:	414
Score:	2158.00	Matches:	414
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.86%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-11 (1-1782) x US-09-211-542A-6 (1-414)

```

QY 538 ATTGCTGACCTGAGTGGCTGTGACGCGTCCATCTGTCAACAAGCCAAATTGTTGTT 597
Db 1 IleaaglyProglutThrleuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTGTGCTGACTGTGTGAGGGAAGAAATCAAGACATTAGTCTTAATTTTGAATC 657
Db 21 SerValaIaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTAACATCAAGACCTCTCTCATGAGAGAAAGACAGACAGCGTTAAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 60
QY 718 AAACCATTTGCTACTGATGTGCGCGGATGTCAACATTAACCTTGAGAAAGCTGACTACTA 777
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
QY 778 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACAGACGACTTCTTGAGTGC 837
Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspPyrPheGluVal 100
QY 838 ATTGATTTTGCAGCGGATGCAACCATTTACTGATCGAACGCGAGGCTTACTTGTCTGAC 897
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValLysPheAlaAsp 120
QY 898 AAAGATGCTGCGCTAACCTTGCGGACCCGAACTGTCCAGAAATTTTGGTAAAGCGGACAT 957
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
QY 958 GTGCGCGTTAGACATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTC 1017
Db 141 ValaGValaIaGProGlyLysGlnLysProIleGlnAsnGlnAlaLysSerValaIaPVal 160
QY 1018 GAATTAAGTGTACAGTTTACCTCTTAAACCTGTGATGACGATTTCAGACCGGCTGCAAA 1077
Db 161 GluLysThrValaIaGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 1078 GATTAAGTGTATTTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTAACTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
QY 1138 GCTCAAGACAAGACATTTTAAACAAAACCAACCCGAGGCTATTCGATTTATGAACGTGAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyLysThrIleLysGlnLysAsp 220
QY 1198 TCTTCAATGTCACATCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
Db 221 SerSerIleValaIaThrHisIleAspAsnAspIlePheArgThrIleLeuProMetAspGlnGly 240
QY 1258 TTTACTTACCGGTGTTAAAAATGGGAACAAGCTTATAGATCAATAAAAATCTGGTCTG 1317

```

```

Db 241 PheThrTyraArgValLysAsnArgGlnGlnAlaIaThrArgIleAsnLysLysSerGlyLeu 260
QY 1318 AATGAGAAATTAACAACACTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGGG 1377
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyraValLeuLysLysGly 280
QY 1378 GAAAAGCCGATATGATCCCTTGAATCGGACGACCTTGAACGTTACCATTAATGATGTT 1437
Db 281 GluLysProTyraSerProPheAspArgSerHisLeuLysLeuPheThrIleLysTyraVal 300
QY 1438 GATGCTGATACCAACGATTTGCTTAAAGTAGAGAGCTCTTAAACGCTAGCGAAGTAAAC 1497
Db 301 AspValaIaSerThrAsnGlnLysLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysAsn 320
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGTGATTAAGCTTAACTTACCAACATCTC 1557
Db 321 LeuAspPheArgAspLeuTyraAspProArgAspLysAlaLysLeuLeuTyraAsnAsnLeu 340
QY 1558 GATGCTTTGGTATTTATGACCTATACCTTAACTGGAAGATGAGATTAACAGATGAC 1617
Db 341 AspAlaPheGlyIleMetAspTyraThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1618 ACCAACCGTATCAATTAACCGTTTATATGAGGCAAGCCGAAAGAGAGAAATGCTAGCTAT 1677
Db 361 ThrAsnArgIleIleThrValLysMetGlyLysArgProGlnGlnLysAlaSerTyra 380
QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAAGAAACGAGAGATTACAGCTACTG 1737
Db 381 HisLeuAlaTyraAspLysAspArgTyraThrGlnGlnGlnLysGlnValLysSerTyraLeu 400
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTACGACGAA 1779
Db 401 ArgTyraThrGlyThrProIleProAspAsnProAsnAspLys 414

```

RESULT 4
 5240845-1
 Patent No. 5240845
 APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;
 ; MUKIWA, BIJUN; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
 ; TSUTOKU
 TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS
 NUMBER OF SEQUENCES: 65
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/549,049
 FILING DATE: 06-JUL-1990
 SEQ ID NO:1:
 LENGTH: 414
 5240845-1

Alignment Scores:
 Pred. No.: 3,87e-210 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 6 Gaps: 0

US-09-940-235-11 (1-1782) x 5240845-1 (1-414)

```

QY 538 ATTGCTGACCTGAGTGGCTGTGACGCGTCCATCTGTCAACAAGCCAAATTGTTGTT 597
Db 1 IleaaglyProglutThrleuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTGTGCTGACTGTGTGAGGGAAGAAATCAAGACATTAGTCTTAATTTTGAATC 657
Db 21 SerValaIaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTAACATCAAGACCTCTCTCATGAGAGAAAGACAGACAGCGTTAAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 60
QY 718 AAACCATTTGCTACTGATGTGCGCGGATGTCAACATTAACCTTGAGAAAGCTGACTACTA 777

```

```

Db      ||| 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
Qy      ||| 778 AAGGCTATCAAGAAACAATGATCGCTAACGTCACAGTAACGACGCTACTTTGAGGTC 837
Db      ||| 81 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspPheGluVal 100
Qy      ||| 838 ATTGATTTTGCAAGCGATGCAACCATTAAGTGAAGCGGACGAGTCTACTTGTGTCAC 897
Db      ||| 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
Qy      ||| 898 AAAGATGTTGGGTAACTTCCGACCCAACTGTCGCAAGATTTTGTCTAACCGACAT 957
Db      ||| 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnIleLeuLeuSerGlyHis 140
Qy      ||| 958 GTGCGCGTTAGACCATATAAGAAACCAATACAAACCAACGAAATCTGTGATGTC 1017
Db      ||| 141 ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy      ||| 1018 GAATATTAAGTGTACAGTTTACTCTCCCTTAAACCTGATGACGATTTGACAGCGGTCTCAA 1077
Db      ||| 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
Qy      ||| 1078 GATACTTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACTCTCAAGATTACTA 1137
Db      ||| 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleLeu 200
Qy      ||| 1138 GCTCAAGCAACAAGCTTTTAAACAAACCAACCCGCGTATACGATTTATGAAGTGCAC 1197
Db      ||| 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluAsp 220
Qy      ||| 1198 TCCCTCAATCGTCACTCATGACATGATCATTTCCGTAAGCTTTTACCAATGATCAAG 1257
Db      ||| 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
Qy      ||| 1258 TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTTATAGATCAATAAAAATCGTCTG 1317
Db      ||| 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
Qy      ||| 1318 AATGAAGAATTAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1377
Db      ||| 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrValLeuLysGly 280
Qy      ||| 1378 GAAGAAGCGTATGATCCCTTATGCGATCACTTGAACCTGTCACATCAATACGT 1437
Db      ||| 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
Qy      ||| 1438 GATGTCGATACCAAGAAATTGCTAAAAAAGTGACAGCTCTTAAACAGCTAGCGAAAGTAA 1497
Db      ||| 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 320
Qy      ||| 1498 TTAGACTTCAGAGATTTATAGATCTCGTGAATAGGCTTAACTACTTACCAAACTTC 1557
Db      ||| 321 LeuAspPheAspArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
Qy      ||| 1558 GATGCTTTTGGTATTTATGACATTAATCTTAACTGAAAAGTGAAGATTAATCAAGTAC 1617
Db      ||| 341 AspAlaPheGlyIleMetAspPheTyrThrLeuThrGlyValGlnAspAsnHisAspAsp 360
Qy      ||| 1618 ACCAAGCGTATACATACCGTTTATATGGGCAAGCGACCGGAGAGAGAGATGTTACTCT 1677
Db      ||| 361 ThrAsnAlaGlyIleThrValTyrMetGlyLysArgProGlnGlnGlnAsnAlaSerTyr 380
Qy      ||| 1678 CATTAGCTTATGATTAAGATGTTTATACCGAAGAAAGAAAGAAATTATACAGTCTAC 1737
Db      ||| 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnValLysSerTyrLeu 400
Qy      ||| 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAAACGACAAA 1779
Db      ||| 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414

```

```

; Sequence 52, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: INENDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (ERO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEO ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-52

Alignment Scores:
Pred. No.: 3,99e-210 Length: 440
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x US-08-560-098A-52 (1-440)
Qy      ||| 538 ATTGCTGACGACCTGACGCTGCTGACGCTCAATCTGTCACACAGCAATGTTGTT 597
Db      ||| 27 IleAlaGlyProGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
Qy      ||| 598 AGCGTTCGTGTAAGTCTGTTGAGGGGACCAATCAAGACATTAAGTCTTAAATTTTGAATC 657
Db      ||| 47 SerValAlaGlyThrValGlnGlnGlnThrAsnGlnAspIleSerLeuLysPheGlnIle 66
Qy      ||| 658 GATCTAATACACGACCTGCTCATGAGGAAGAAGACAGACGAAGGCTTAAGTCCAAAATCA 717
Db      ||| 67 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlnGlnLysSerProLysSer 86
Qy      ||| 718 AAACCATTTGCTACTGATTAAGTGGCGGATGTCACATTAACCTTGAAGAAAGTGAATCTA 777
Db      ||| 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
Qy      ||| 778 AAGGCTATTAAGAAACAATGATCGCTTACGTCACAGTAACGACGACTTGAAGTCT 837
Db      ||| 107 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspPheTyrPheGlnVal 126

```

```

QY 838 ATTGATTTGCAAGGATGCAACCATTAATGATCGAAGGCAAGGCTTACTTGGCTGAC 897
Db 127 ILeaspPheAlaSerValAspAlaThrIleThrAspArgAsnGlyValValYrPheAlaAsp 146
QY 898 AAAAGATGTTGGTAACTTGGCCGACCCCACTGTCGAAGATTGCTTAAGCGGACAT 957
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 166
QY 958 GTGGCGGTTAGACCATATTAAGAAAAACCAATATCAAAACCAAGCGGAATCTGTGATGG 1017
Db 167 ValArgValArgProYrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATATCTGACATTAATCTCCCTTAACCCCTGATGACATTTGACAGAGCTCTCAAA 1077
Db 187 GluYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1078 GATACATACTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 1137
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
QY 1138 GCTCAAGCAAAAGATTTTAAACAAAACCAAGCGGTATACGATTTATGAAGCTGAC 1197
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyYrThrIleYrGluArgAsp 246
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTAACATTTTACCAATGATCAAGAG 1257
Db 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1258 TTTAATTAACCGTGTAAAAATCGGAGACAGCTTATAGCATCAATAAAAATCTGGTCTG 1317
Db 267 PheThrYrArgValLysAsnArgGluGlnAlaYrArgIleAsnLysSerGlyLeu 286
QY 1318 AATGAAGAAAAATTAACAACATGACCTGATCTGTGAAAAATATTAAGCTCTTAAAAAGG 1377
Db 287 AsnGluGluIleAsnAsnThrAspLeuLysSerGluYrYrValLeuLysGly 306
QY 1378 GAAAGCCGATGATGATCCCTTGTGATCGACATCTTGAAGCTGTCACCATCAATACGTT 1437
Db 307 GluLysProYrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysYrVal 326
QY 1438 GATTCGATATCAACCAAGATGCTTAAAAAAGTGAAGAGCTTAAAGCTAAGCAAGTAC 1497
Db 327 AspAlaAspThrAsnGluLeuLysSerGluGlnLeuThrAlaSerGluArgAsn 346
QY 1498 TTAACTTCAAGATTTTATGATCTCTGTGATTAAGGCTAAACTTACTTACCAATCTTC 1557
Db 347 LeuAspPheArgAspLeuYrAspProArgAspLysAlaLysLeuLysYrAsnAsnLeu 366
QY 1558 GATGCTTTTGGTATTAATGACATATCTTAATCTGAAAAAGTAGAGATTAATCAGATGAC 1617
Db 367 AspAlaPheGlyIleMetAspYrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY 1618 ACCAACGATATCATTAACGTTTATATGAGCAAGCAAGCCGAAAGAGAGATCTAGCTAT 1677
Db 387 ThrAsnArgIleIleThrValIleMetGlyLysArgProGluGluGluAsnHisAspYr 406
QY 1678 CATTAGCTTATGATTAAGATCGTTTATACCGAAGAAAGAAAGATTATACGATCTG 1737
Db 407 HisLeuAlaYrAspLysAspArgYrThrGluGluGluArgGluValIleYrSerYrLeu 426
QY 1738 CGTATATCAGGAGACCTTATACCTGATTAACCTTAACGACAA 1779
Db 427 ArgYrThrGlyThrProIleProAspAsnProAsnAspLys 440

```

```

; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:
Pred. No.: 1,24e-209 Length: 413
Score: 2153.00 Matches: 413
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.70% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x US-10-360-101-264 (1-413)
QY 538 ATTGCTGACCTGATGCGTGTGACCCGTCATCTGTCAACAACAGCCAAATGTTGTT 597
Db 1 ILeAlGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
QY 598 AGCGTGTCTGTATCTGTGAGGAGCAAGATCAAGACATTAAGTCTTAAATTTTGAATC 657
Db 21 SerValAlaGlyThrValGluGluYrThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 658 GATCTTAACATCAGACCTGCTCATGAGCAAGAAAGCAAGAGCTTAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGluLysSerProLysSer 60
QY 718 AAACCATTTGCTGATGATGTGGCGGATGTCATATACTTGAAAGCGTCTTACTA 777
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY 778 AAGGCTATTCAGAACCAATGATCGCTAACGTCACAGTAAACGACATCTTGAAGTCT 837
Db 81 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerHisAspAspYrPheGluVal 100
QY 838 ATTGATTTGCAAGCAATGCAACCATTAATCTGATGAAACGCAAGCTTACTTGTCTGAC 897
Db 101 ILeaspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValYrPheAlaAsp 120
QY 898 AAAAGATGTTGGTAACTTGGCCGACCCCACTGTCGAAGATTGCTTAAGCGGACAT 957
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
QY 958 GTGGCGGTTAGACCATATTAAGAAAAACCAATCAAAACCAAGCGGAATCTGTGATGG 1017
Db 141 ValArgValArgProYrLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATATCTGATACGTTTATCTCCCTTAACCCCTGATGACATTTGACAGAGCTCTCAA 1077
Db 161 GluYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATACATACTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1138 GCTCAAGCAAAAGATTTTAAACAAAACCAAGCGGTATACGATTTTATGAAGCTGAC 1197
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyYrThrIleYrGluArgAsp 220
QY 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCGTAACATTTTACCAATGATCAAGAG 1257
Db 221 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1258 TTTAATTAACCGTGTAAAAATCGGAGCAAGCTTATAGATCAATAAAAATCTGGTCTG 1317

```

```

RESULT 6
US-10-360-101-264
; Sequence 264, Application US/10360101
; Patent No. 6861236
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673

```


Qy	1387	TATGATCCCTTGAATCGAGACACTTGAACTGTTCAACATCAATACGTTGAGTGCAT	1446
Db	664	TyrAspProPheAspArgSerHisIleuysIeuPheTrIleYsYrValAspValAsp	683
Qy	1447	ACCAACGAATGCTAAAAAGTAGAGCGCTTTAAACAGCTAGCGAAGCTTAACCTTGAAGCTTC	1506
Db	684	ThrAsnGluIleuIeuIeuYseGluGlnIleuIeuThrIAsSerGluWAGAsnIeuAspHe	703
Qy	1507	AGAGATTATACGATCCTCGTGATTAAGCTTAACCTACTCTACAACAATCTCGATGCTTTT	1566
Db	704	ArgAspIeuYrAspProArgAspIysAlaIysIeuIeuYrAsnAsnIeuAspAlaPhe	723
Qy	1567	GGTATTATAGACTATACCTTAACCTGAAAGATAGAGATATATCAAGATGACCAACCGT	1626
Db	724	GlyIleuIeuAspYrThrIeuThrGlyIysValGluAspAsnHisAspAspThrAsnArg	743
Qy	1627	ATCATTAACCGTTTATATGAGCAACGCAACCCGAAAGAGAGAAATGCTAGCTATCATTTAGCC	1686
Db	744	IleIleuThrValIyrMetGlyIysArgProIuGluGluAsnAlaSerYrHisIeuAla	763
Qy	1687	TATGATTAAGATCGTTATATCGAAGAGAAGAGAAAGTTTACAGCTACCTGCGTTATACA	1746
Db	764	TyrAspIysAspArgYrThrGluGluGluArgIuValYrSerYrIeuArgYrThr	783
Qy	1747	GGGACACCTATACCTGATTAACCTTAACGACAAA	1779
Db	784	GlyThrProIleProAspAsnProAsnAspIys	794
RESULT 8			
US-07-854-596B-28			
Sequence 28, Application US/07854596B			
Patent No. 5434073			
GENERAL INFORMATION:			
APPLICANT: Dawson, Keith M			
APPLICANT: Hunter, Michael G			
APPLICANT: Czaplowski, Lloyd G			
TITLE OF INVENTION: Proteins and nucleic acids			
NUMBER OF SEQUENCES: 73			
CORRESPONDENCE ADDRESS:			
ADDRESSER: Dr. John J. McDonnell			
STREET: Ten South Wacker Drive, Suite 3000			
CITY: Chicago			
STATE: IL			
COUNTRY: USA			
ZIP: 60606			
COMPUTER READABLE FORM:			
MEDIUM TYPE: floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/07/854,596B			
FILING DATE: 03-JUN-1992			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: McDonnell, John J			
REGISTRATION NUMBER: 26,949			
REFERENCE/DOCKET NUMBER: 92,337			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 312-715-1000			
TELEFAX: 312-715-1234			
TELEX: 910-221-5317			
INFORMATION FOR SEQ ID NO: 28:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 499 amino acids			
TYPE: amino acid			
TOPOLOGY: linear			
MOLECULE TYPE: protein			
US-07-854-596B-28			
Alignment Scores:			
Pred. NO.: 1.42e-208			
Score: 2143.00			
Length: 499			
Matches: 413			

Percent Similarity:	96.29%	Conservative:	2
Best Local Similarity:	95.82%	Mismatch:	16
Query Match:	68.38%	Indels:	0
DB:	1	Gaps:	0
US-09-940-235-11 (1-1782) x US-07-854-5968-28 (1-499)			
QY 487	ACCTCTGCGACGACCACTCGAGCGGATCTGGCCCTTACCCGATGTTCTGATTGCTGGA	546	
DB 69	ThrlleaserlleaialaIyaglugluglValSerleuambyargllealagly	88	
QY 547	CCCTAGAGGCGCTGACGCGCTCATCTGTCAACAAACGCAATGGTGTAGCGTCT	606	
DB 89	ProclutrlpeuleuamPargProSeValasmnserglneValaSerVala	108	
QY 607	GGTACTGTGGAGGAGCAATCAAGACATTAGTCTTAATTTTTTGAATCGATCTAAC	666	
DB 109	glYlthralgluglYlthraenglnasrlleSerleuysPhepeglnleasPheuthr	128	
QY 667	TCACGACCTGCTCATGAGGAAGACAGAGCGTTAAGTCCAAATCAAAACATTT	726	
DB 129	SezarProclamlhlsrglyglYlYlthrglnglnglYleuSerProlysSerlyProphe	148	
QY 727	GCTACTGATAGTGCGGAGTGCTCATTAATCTGAGAAAGTGACTTAAGGCTATT	786	
DB 149	AlathrasPserglYalameCProhlslsYleuclulysalaePleuYalale	168	
QY 787	CANAGCAAAATTGATCGCTAACGTCACAGTAACAGCATCTTGTGAGTCATTGATTTT	846	
DB 169	glnglnglnglnleuilealaaenValHlsSerlasmPserlyrPhegluValleasPhe	188	
QY 847	GCAAGCAGATGCACCATTTACTGATGAAACGGCAAGCTTACTTGTCTGACAAAGATGT	906	
DB 189	AlaseraplatlthrllethrasPargasnclYlYlValTyPhealaaPlyasPclY	208	
QY 907	TCCGTAACTTTGCCGACCCCACTGTCGCAAAATTTTTCGAACCGCACATGCGCGCTT	966	
DB 209	SerValThrluProltnlgnProValglnglnglnPheleuSerglYHlsvalaTgVal	228	
QY 967	AGACCATATTAAGAAAGAAACCAATCAAAACCAACGCAATCTGTGATGTGAATATCT	1026	
DB 229	AlrgProlyrlysglYlYlProlleglhnsnglnalalyServalasPvalglulYrThr	248	
QY 1027	GTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGGTCTCAAGATCTAAG	1086	
DB 249	ValglInPhehrProlenPasnProasParsPheargProglYleuYasPrlhYls	268	
QY 1087	CTATTGAAACACTAGCTATGGGTGACACCATCTCAATCTCAAGAAATTCTAGCTCAAGA	1146	
DB 269	leuYeuYsThrlleualaleglYasPThrllethSerglnglnglnleuValaglnala	288	
QY 1147	CAAGGCAATTTTAAACAAAAACCAACCGGCGTATACGATTTATGAAACGGACTCCGCAATC	1206	
DB 289	GlmserrlleuasnlyPthrlasPocglYrYrThrlleTyrglnlYlYlYlYlYlYlYlYlYl	308	
QY 1207	GTCACTGATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGATTTACTAC	1266	
DB 309	ValthrlhlsasPasnasPrllePheargThrlleuPPrometasPnglnglnPhehrTyx	328	
QY 1267	CGTGTAAAAATCGGGAACAAAGCTTATAGGATCAATAAAAATCTGGCTGATGAAAGA	1326	
DB 329	HlsVallylsasPasnaglnglnglnalalyrclulleasnlylYlYlYlYlYlYlYlYlYl	348	
QY 1327	ATAAACAAACCTGACCTGATCTCGAGAAATATACGCTTAAAAAAAGGGAAGACGCG	1386	
DB 349	IlleasnenthlasPleulleSerglnlysrTyxValleuYlYlYlYlYlYlYlYlYlYlYlYlYl	368	
QY 1387	TATGATCCCTTGTGATGCGAGTCACTTGAACTGTTCACCATCAATACGTTGATGTGAT	1446	
DB 369	TyrAsPrlorPheasPargserHlsleuYlsleuPhehrThrlleYlYlYlYlYlYlYlYlYl	388	
QY 1447	ACCAAGCAATTTGCTAAAAAGTAGACAGCTCTTAAACGCTAGCGAAAGTAACTTACATTC	1506	

Db 389 ThrAnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAnLeuLeuAspPhe 408
 QY 1507 AGAGATTATACGATCTCGTGATTAAGCTTAACCTACTACAAACAATCTCGATGCTTTT 1566
 Db 409 ArgAspLeuTyrAspProArgAspLysAlaIleLysLeuLeuTyrAsnAsnLeuAspAlaPhe 428
 QY 1567 GGTATTATGACATTAACCTTAACCTGAAAGTAAGAGATTAATCAACGATGACACCAACCGT 1626
 Db 429 GlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAspThrAsnArg 448
 QY 1627 ATCATTAACCGTTTATATGGGCAAGCAACCCGAAAGAGAAATGCTAGCTTACATTAGCC 1686
 Db 449 IleIleThrValTyrIleGlyLysArgProGluGlnGlnAsnAlaSerTyrHisLeuAla 468
 QY 1687 TATGATAAAGATCGTTATATCCGAAAGAAAGAGATTATACAGCTTACCTGCGTTATACA 1746
 Db 469 TyrAspLysAspArgTyrThrGlnGlnGlnArgGlnValTyrSerTyrLeuArgTyrThr 488
 QY 1747 GGGACACCTATACCTGATTAACCTTAACGACAAA 1779
 Db 489 GlyThrProIleProAspAsnProAsnAspLys 499

RESULT 9

US-08-488-940-3
 ; Sequence 3, Application US/08488940
 ; Patent No. 5854049

; GENERAL INFORMATION:
 ; APPLICANT: Reed, Guy L.
 ; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,940
 ; FILING DATE: 09-JUN-1995
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Frazer, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 05433/009001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 813 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-488-940-3

Alignment Scores:

Pred. No.: 3,266-208 Length: 813
 Score: 2140.50 Matches: 413
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 0
 Query Match: 68.30% Indels: 1
 Gaps: 1

US-09-940-235-11 (1-1782) x US-08-488-940-3 (1-813)

QY 538 ATTGCTGACCTGAGTGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGGTGTG 597
 Db 1 IleIleIleGlyProGluTutThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 598 AGCGTGTGCTGATCTGTGAGGGGACGAATCAAGACATTAGCTTTAAATTTTGAATC 657
 Db 21 SerValIleGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 658 GATTCATCATGACGACCGCTCATGGAGGAAAGACAGGCAAGGCTTAAGTCCAAATCA 717
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 60
 QY 718 AACCAATTTGCTAGTAGTAGGCGCGATGACATAAATCTTGAGAAAGCTGACTTACTA 777
 Db 61 LysProIleAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
 QY 778 AAGCTATTCAAGAACCAATTGATGCTTACGTCACAGTTAACGACGACTTGTGAGGTC 837
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGluVal 100
 QY 838 ATTGATTGCAAGCGATGCAACCATTAATGATGGAACCGGCAAGCTTACTTGTCTAC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 898 AAGATGATGTCGGTAACCTTGCCGACCCCAACCTGTCGAAGAATTTTGTGAAGCGACAT 957
 Db 121 LysAspLysSerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 958 GTGGCGTTGACCATTAAGAAAGAAACCAATACAAACCAAGGAAATCTGTGATGTG 1017
 Db 141 ValArgValArg--TyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 159
 QY 1018 GAATTAATGTGATGATTATCTCCCTTAAACCTGATGACATTTCAGACCGGTCTCAA 1077
 Db 160 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
 QY 1078 GATCTAAGCATTAAGAAACCTAGCTATGGTGACACCATCAACCTCAAGAAATTTCTA 1137
 Db 180 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 199
 QY 1138 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCAAGGCTTATGACATTATGACAGTAC 1197
 Db 200 AlaGlnIleGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 219
 QY 1198 TCCTCAATCGTCACTGACATGACATGACATTTCCGTACGATTTTACCAATGATCAAGAG 1257
 Db 220 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 239
 QY 1258 TTTAATTAACCGTGTAAATTCGGGAACAAGCTTATAGATCATTAATAAATCTGGTCTG 1317
 Db 240 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 259
 QY 1318 AATGAAGAATTAACCAACACTGACCTGATCTGAGAAATATTAAGTCTTAAAGGAGG 1377
 Db 260 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 279
 QY 1378 GAAAGCCGTATGATCCCTTGTATCGAGTCACTTGAACCTGTTCAACATCAATAGCTT 1437
 Db 280 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 299
 QY 1438 GATGTCGATACCAACGAATGCTTAAAGTAGAGAGCTTTTAAACGACTAGCGAAGCTTAC 1497
 Db 300 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 319
 QY 1498 TTAGACTTCAGAGATTATATCGATCCGCGATGATTAAGGCTAAACATCTCAACAATCTC 1557
 Db 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 339
 QY 1558 GATGCTTTGTGATTAATGACTATACCTTAAGTGAAGAAAGTAAGAGATTAATCAAGATGAC 1617
 Db 340 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 359
 QY 1618 ACCAACCGTATCATTAACCGTTTATATGCGCAAGGACCCGAAGAGAAATGCTAGCTAT 1677

```

Db      360  ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluValSerTyr 379
Qy      1678  CATTAGCCTTATGATAAAGATGCTTATACCGAAGAAGACGAGACTTACCTACTG 1737
Db      380  HisLeuAlaTyrAspLysAspArgTyrThrGluGluValArgGluValTyrSerTyrLeu 399
Qy      1738  CGTTATACAGGAGACCTTATACCTGATACCTTACGACGAA 1779
Db      400  ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 413

RESULT 10
US-07-854-596B-26
; Sequence 26, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-26

Alignment Scores:
Pred. No.: 1,696-207 Length: 415
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-596B-26 (1-415)
Qy      538  ATTGCTGACCTGATGAGCTGCTAGACCTGCATCTGTCAACAACGCAATTGTTGTT 597
Db      2  IleAlaGlyProGluThrPheLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 21
Qy      598  AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGCTTAAATTTTGAATC 657
Db      22  SerValAlaGlyThrValGluGluGlyThrAsnGlnAspIleSerLeuYspHepGluIle 41
Qy      658  GATCAACATGACGACCTGCTCATGAGAGAAAGACGACGACGCTTAAGTCCAAATCA 717

```

```

Db      42  AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 61
Qy      718  AAACCATTTGCTATGATAGTGGCGATGTCACATTAACCTTGAGAAAGCTGACTACTA 777
Db      62  LysProHeaAlaThrAspSerGlyAlaMetProHisLysLeuGluValAspLeuLeu 81
Qy      778  AAGCTATTCCAGAAACATTTGATGCTTACGCTCCACAGTACGACGACGACTTTGAGGCT 837
Db      82  LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 101
Qy      838  ATTGATTTTGGACGACGATGACCAACCATTTACTGATGGAACGGCAAGTCTTGTCTGAC 897
Db      102  IleAspPheAlaSerAspAlaThrIleThrAspArgGlnGlyValTyrPheAlaAsp 121
Qy      898  AAAGATGGTTCGTGTAACCTTGCCGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 957
Db      122  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 141
Qy      958  GTGCGCGTTAGACCATTAAGAAAGAAACCAATACAAACCAAGGGAATCTGTGATGTG 1017
Db      142  ValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 161
Qy      1018  GAATATACGTGACGTTTACTCCCTTAAACCTGATGACGATTTCCAGACGAGTCTCAAA 1077
Db      162  GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 181
Qy      1078  GATATCAAGCTATTGAAAACATGCTATCGGTGACACCATCACTCAAGATTACTA 1137
Db      182  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 201
Qy      1138  GCTCAACGACAAAGCATTTTAAACAAACCAACCGAGCTTATAGATTATGACGTCGAC 1197
Db      202  AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 221
Qy      1198  TCCTCAATGCTCATCTGACATGACATGATTTCCGTGACATTTTACCAATGATCAAG 1257
Db      222  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 241
Qy      1258  TTATCTTACCGGTGTTAAAAATCGGAAACAGCTTATAGATCAATAAAAATCTGCTCTG 1317
Db      242  PheThrTyrHisValLysValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGlyLeu 261
Qy      1318  AATCAAGAATTAACCAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGG 1377
Db      262  AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 281
Qy      1378  GAAAAGCGGTATGATCCCTTGTGATCGGACGATCACTTGAACGTTCAACATCAATACGTT 1437
Db      282  GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 301
Qy      1438  GATGTCGATACCAACGAATTGCTTAAAAAGTAGAGAGCTTTAACAGCTTAGCGAACGTAAAC 1497
Db      302  AspValAsnThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 321
Qy      1498  TTAGACTTTCAGAGATTTTATACGATCCCTCGTGAATAAGGTTAACTTCAACAATCTC 1557
Db      322  LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 341
Qy      1558  GATGCTTTTGGTATTTATGACTATACCTTAACTGGAAGAAAGTAGAGATTAACGATGAC 1617
Db      342  AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 361
Qy      1618  ACCAACGCTATCATTAACGTTTATATGAGGACGAGGACCCGAGAGAGAGAAATGCTAGCTAT 1677
Db      362  ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluValAsnAlaSerTyr 381
Qy      1678  CATTACCTTATGATAAAGATCGTTATACGGAAGAAAGACGAAAGTTTACGCTACTG 1737
Db      382  HisLeuAlaTyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 401
Qy      1738  CGTTATACAGGAGACCTTATACCTGATACCTTACGACGAA 1779
Db      402  ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 415

```

RESULT 11

US-07-854-596B-19

; Sequence 19, Application US/07854596B

; Patent No. 5434073

GENERAL INFORMATION:

; APPLICANT: Dawson, Keith M

; APPLICANT: Hunter, Michael G

; APPLICANT: Czaplowski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. John J. McDonnell

; STREET: Ten South Wacker Drive, Suite 3000

; CITY: Chicago

; STATE: IL

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/854,596B

; FILING DATE: 03-JUN-1992

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: McDonnell, John J

; REGISTRATION NUMBER: 26,949

; REFERENCE/DOCKET NUMBER: 92,337

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-715-1000

; TELEFAX: 312-715-1234

; TELEX: 910-221-5317

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 435 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-07-854-596B-19

Alignment Scores:

Pred. No.: 1,73e-207

Score: 2132.00

Percent Similarity: 99.03%

Best Local Similarity: 98.79%

Query Match: 68.03%

DB: 1

US-09-940-235-11 (1-1782) x US-07-854-596B-19 (1-435)

Length: 435

Matches: 409

Conservative: 1

Mismatches: 4

Indels: 0

Gaps: 0

QY 538 ATTGCTGACCTGAGTGGCTGCTAGACGCTCATCTGTCAACAACGCAATTGTTGTT 597

DB 22 Ileaaglyprrcglutrrpleuemaapargprosevalashnanserglnleuval 41

QY 598 AGCGTGTGCTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 657

DB 42 SerValaIaaglyThrValgluglyThrAsnGlnAspIleSerleuysPhepegluile 61

QY 658 GATCTAACATCAAGCACTCTCTCATGGAGAAAGACAGACAGAGCTTAAGTCCAAATCA 717

DB 62 AspLeuThrSerArgProAlaHisgIyIyIySThGlnGlnGlyleuSerProlysSer 81

QY 718 AAACCATTTGCTACTGATGTGGCGGAGTGTCACTTAACCTTGAAGAAGCTGACTACTA 777

DB 82 LysPrrPrrAlaThrAspSerGlyAlaMetProHisLysleuGlyLysAlaAspLeuLeu 101

QY 778 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAAGACAGACTTGTAGAGTC 837

DB 102 LysAlaIleGlnGlnGlnleuIleAlaAsnValHisSerAsnAspAspIyrrPhegluVal 121

QY 838 ATTGATTTTGCAGACGATGCAACCAATTACTGATCGAAAGCGCAAGTCTACTTTGCTGAC 897

DB 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIyrrPheAlaAsp 141

QY 898 AAAAGATGGTTGGTAACTTCCGACCCAGACCTGTCCAGAAATTTTGTGAAGGCAAT 957

DB 142 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheleuSerGlyHis 161

QY 958 GTGGCGCTTGGACCATTTAAAGAAACCAATACAAACCAAGGCAATCTGTGAATGTG 1017

DB 162 ValArgValAlaGProIyrrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 181

QY 1018 GAATATACCTGATGAGTTACTCCCTTAAACCCCTGATGACGATTTACAGACGCTCAAA 1077

DB 182 GluIyrrThrValGlnPheThrProleuAsnProAspAspPheAspProGlyleuLys 201

QY 1078 GATCTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACTTCAAGAAATTACTA 1137

DB 202 AspThrLysleuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnleuLeu 221

QY 1138 GCTCAAGCACAAAGCAATTTTAAACAAACCAAGGCTTACGATTTTAAAGACGTAC 1197

DB 222 AlaGlnAlaGlnSerIleleuAsnLysThrIleAspGlyIyrrThrIleIyrrGluArgAsp 241

QY 1198 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257

DB 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleleuProMetAspGlnIle 261

QY 1258 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTTAAGATCAATAAAAATCTGCTGTG 1317

DB 262 PheThrIyrrHisValLysValAsnArgGlnGlnAlaIyrrGluIleAsnLysSerGlyleu 281

QY 1318 AATGAAGAATTAACAACACTGACCTGATCTGTGAGAAATTTTCCTCTTAAAAAGGG 1377

DB 282 AsnGlnGlnIleIleAsnAsnThrAspLeuIleSerGluLysIyrrIyrrValLeuLysIleGly 301

QY 1378 GAAAGCCGTATGATCCCTTGTATCGAGTCACTTGAACGTTCAACATTAATAGTT 1437

DB 302 GluLysProIyrrAspProPheAspArgSerHisleuLysleuPheThrIleLysIyrrVal 321

QY 1438 GATGTCGATCAACAAGATTTGCTAAAAAGTAGAGGCTTTAAGCTAGGAGCAATGTAAC 1497

DB 322 AspValAsnThrAsnGlnleuLeuLysSerGlnleuLeuThrAlaSerGluArgAsn 341

QY 1498 TTAGACTTCAAGATTTATACGATCTCTGTGATTAAGGCTTAATCTTCAACAATCTC 1557

DB 342 LeuAspPheArgAspLeuIyrrAspProArgAspLysAlaLysleuLysIyrrAsnAsnleu 361

QY 1558 GATGCTTTTGTATTAAGACTATACCTTAACCTGGAAGAGTAAATCAAGATGAC 1617

DB 362 AspAlaPheGlyIleMetAspIyrrThrLeuThrGlyLysValGluAspAsnHisAspAsp 381

QY 1618 ACCAAGCGTATCAATTAACGTTTATATGGGAAGGACCCGGAAGAGAGAAATGCTAGCTAT 1677

DB 382 ThrAsnArgIleIleThrValIyrrMetGlyLysArgProGlnGlnLysAlaSerIyrr 401

QY 1678 CATTAGCTTATGATTAAGATCGTTATACGAGAAAGACAGAGATTTACGCTACTG 1737

DB 402 HisleuAlaIyrrAspLysAspArgIyrrThrGlnGlnGlnLysValIyrrSerIyrrleu 421

QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAA 1779

DB 422 ArgIyrrThrGlyThrProIleProAspAsnProAsnAspLys 435

RESULT 12

US-07-854-596B-15

; Sequence 15, Application US/07854596B

; Patent No. 5434073

GENERAL INFORMATION:

; APPLICANT: Dawson, Keith M

; APPLICANT: Hunter, Michael G

; APPLICANT: Czaplowski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids

```

; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-596B-15

Alignment Scores:
Pred. No.: 1.74e-207 Length: 440
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
DB: Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-596B-15 (1-440)
QY 538 ATGCTGACCTGAGTGGCTGTAGACCTCATCTGTAAACAAGCCAAATTGCTGTT 597
DB 27 TTTGCTGAGTGGCTGTAGACCTCATCTGTAAACAAGCCAAATTGCTGTT 597
QY 598 AGCGTTGCTGTAAGTGTGAGGGAAGCAATCAAGACATTAGTCTTAAATTTTGAATC 657
DB 47 SerValAlaGlyThrValGluGlyThrValGlnAspIleSerLeuLysPheGluIle 66
QY 658 GATTTACATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyLysThrGlnGlnGlyLeuSerProLysSer 86
QY 718 AAACATTGCTACTGATGATGCGCGGATGTCAATAACTTGAAAGCTGACTTACTA 777
DB 87 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 106
QY 778 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAAGACGACTTGTGAGTC 837
DB 107 LysAlaIleGlnGlnGlnLeuIleAlaSerValHisSerAsnAspAspLysPheGlnVal 126
QY 838 ATGATTTTGAAGGATGACCAACATTAAGTCAAGCAAGGAGGCTTACTTGTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 898 AAAAGTGTGGTGAAGTGTGACCTGCGACCACTGTGCAAGAAATTTTGTAAAGCGACAT 957
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnLysPheLeuSerGlyHis 166
QY 958 GTGGCGCTTAGCACAATATTAAGAAACCAATCAAAACCAAGCGAAATCTGTGATG 1017

```

```

DB 167 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATACGTATACAGTTTACTCCCTTAAACCGTAAACGATTTCAACAGGCTGTA 1077
DB 187 GlnTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
QY 1078 GATATCAAGGCTATTTGAAACACATAGCTATCGGTGACACCATCAATCATGCAAGATTACTA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 226
QY 1138 GCTCAAGCACAAGCAATTTTAAACAAACACACCGAGCTTACGATTTTGAACGTGAC 1197
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnLysArg 246
QY 1198 TCTGCAATGCTCACTGATGACATGACATTTTCCGTCACATTTTACCAATGATCAAGAG 1257
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1258 TTTACTTACCGGTGTTAAATCGGGAACAGCTTATAGATCAATATAAAATCGTCTG 1317
DB 267 PheThrTyrHisValLysAsnArgGlnAlaIleTyrGlnLysLysSerGlyLeu 286
QY 1318 AATGAAGAAATTAACAACACTGACCTGATCTGAGAAATATTACGTCCTTAAAGAGG 1377
DB 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
QY 1378 GAAAAGCGGTATATATCCCTTGTATCGGAGTCACTTGAACCTGTTCACATCAATACGTT 1437
DB 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGTCGATCAACCAAGATTGTCTAAAGAGACAGCTCTTACAGCTAGCAAGACGTAAC 1497
DB 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 346
QY 1498 TTAGCTTCAAGATTTATACGATCTCGTATAGGCTTAACTACTTCAACAAATCTC 1557
DB 347 LeuAspPheArgAspLeuTyrAspProAlaArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1558 GATGCTTTTGGTATTTATGACATATACCTTACTGAGAAAGTAAAGATTAATCACGATGAC 1617
DB 367 AspAlaPheGlyLysLeuAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 386
QY 1618 ACCAACCGTATCAATACCGTTTATGCGGACGACCGCAAGAGAGAAATGCTAGCTAT 1677
DB 387 ThrAsnArgLysIleThrValTyrMetGlyLysArgProGlnGlnGlnAlaSerTyr 406
QY 1678 CATTTAGCCTATGATTAAGATCGTATATACCGAAGAGAAAGAGATTTACAGCTACTG 1737
DB 407 HisLeuAlaIleTyrAspLysAspArgTyrThrGlnGlnGlnLysGlnValTyrSerTyrLeu 426
QY 1738 CGTTATCAAGGACACCTTATACCTGATTAACCTTAACGACAA 1779
DB 427 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440

RESULT 13
US-07-854-596B-47
; Sequence 47, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplewski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B
 FILING DATE: 03-JUN-1992
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949
 REFERENCE/DOCKET NUMBER: 92,337
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 910-221-5317
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 483 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-854-596B-47

Alignment Scores:
 Pred. No.: 1,83e-207 Length: 483
 Score: 2132.00 Matches: 409
 Percent Similarity: 99.03% Conservative: 1
 Best Local Similarity: 98.79% Mismatches: 4
 Query Match: 68.03% Indels: 0
 DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-596B-47 (1-483)

QY 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACAACACCAATTGGTTGT 597
 DB 1 TLeaAgLyProGluTrpLeuLeuAspArgProSerValAsnSeriGlnLeuValVal 20
 QY 598 AGCGTGTGCTGCTGCTGTTGAGGGAACCAATCAAGACATTAGTCTTAAATTTTGAATC 657
 DB 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 658 GATCTAACATCAGACCTGCTCATGAGGAAGAAGAGCAAGCTTAAGTCCAAATCA 717
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlyLeuSerProLysSer 60
 QY 718 AAACCTTTGCTACTGATAGTGGCGGAGATGTCATATAATTGAGAAAGCTGACTTACTA 777
 DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 80
 QY 778 AAGGCTATTCAAGAAATTTGATCGTACGCTAAGCAAGCACTACTTTGAGGTC 837
 DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgLysPheGlnVal 100
 QY 838 ATGTATTTTGAACGCGATGCAACCATTAAGTGAACGCGAAGTCTACTTTGCTGAC 897
 DB 101 IleAspPheAlaSerSerAspAlaThrIleThrAspArgAsnGlyLysValLysPheAlaAsp 120
 QY 898 AAAAGATGTTCCGTTAACTTGGCCGACCCAACTGCTCCAGAAATTTTGTAAACGACAT 957
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnLysLeuLeuSerGlyHis 140
 QY 958 GTGCGCGTTAGACCAATATAAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTC 1017
 DB 141 ValArgValArgProLysGlyLysProLysGlnAsnGlnAlaLysSerValAspVal 160
 QY 1018 GAATACTAGTACAGTTTACCTCTTAAACCTGATGACGATTTGACAGCGGTCTGAAA 1077
 DB 161 GlnLysThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 QY 1078 GATACTAACCTATTGAAAACACTAGTATCGGTGACACCATCTCAAGATTACTA 1137
 DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200

QY 1138 GCTCAGACAAAGACATTTTAAACAAAACCAACCCAGGCTATGACATTTATGACGTGAC 1197
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyLysThrIleLysGlnArgAsp 220
 QY 1198 TCCTCAATCGTCACCTGACGACAAATGACATTTTCCGTAGATTTTACCAATGATCAAGG 1257
 DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
 QY 1258 TTTACTTACCGGTGTTAAATAATCGGACACAGCTTTAGATCAATATAAAATCTGCTCG 1317
 DB 241 PheThrLysHisValLysAsnArgGlnGlnAlaLysGlnLysLeuLysSerLysLeu 260
 QY 1318 AATGAAGAAATTAACACACTGACCTGATCTCTGAGAAATATTAAGCTTTAAAGAGG 1377
 DB 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysLysLysValLeuLysGly 280
 QY 1378 GAAAAGCCGTATGATCCTTTGATGCGAGTCACTTGAATCTTTACCAATCAATACGTT 1437
 DB 281 GlnLysProLysArgProPheAspArgSerHisLeuLysLeuPheThrIleLysVal 300
 QY 1438 GATGTCGATCAACCAAGAAATGCTAAAGTAGAGCGCTTATACGCTTAAGCAAGTAC 1497
 DB 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 320
 QY 1498 TTGACCTTCAGAGATTTTATAGATCTCTGATTAAGCTTAAGCTTACTCTACCAACATCTC 1557
 DB 321 LeuAspPheArgAspLeuLysThrAspProArgAspLysAlaLysLeuLysThrAsnLeu 340
 QY 1558 GATGCTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGATATCAAGATGAC 1617
 DB 341 AspAlaPheGlyIleMetAspLysThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
 QY 1618 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAAAGAGAAATGCTATGCTAT 1677
 DB 361 ThrAsnArgIleIleThrValLysMetGlyLysArgProGlnGlnGlnValLysSerLys 380
 QY 1678 CATTTAGCCTATGATTAAGATCGTTATACGAGAAAGAACGAAAGTTTACAGCTACCTG 1737
 DB 381 HisLeuAlaLysAspLysAspArgLysThrGlnGlnGlnGlnValLysSerLysLeu 400
 QY 1738 CGTTATACGAGGACACTATACCTGATTAACCTTACGACAA 1779
 DB 401 ArgLysThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 14
 US-07-854-596B-35
 Sequence 35, Application US/07854596B
 Patent No. 5434073
 GENERAL INFORMATION:
 APPLICANT: Dawson, Keith M
 APPLICANT: Hunter, Michael G
 APPLICANT: Czaplowski, Lloyd G
 TITLE OF INVENTION: Proteins and nucleic acids
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dr. John J. McDonnell
 STREET: Ten South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B
 FILING DATE: 03-JUN-1992
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949


```

REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 859 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-5968-35

Alignment Scores:
Pred. No.: 2,44e-207 Length: 859
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x US-07-854-5968-35 (1-859)
QY 538 ATTGCTGACCTGAGTGGCTGCTAGACGCTCCATCTGTCAACAACAGCCAAATTGTTGTT 597
DB 22 TleaaaglyProglutThrleuleuaspargProseivalashnshserglneuval 41
QY 598 AGCGTGTGCTACTGTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 657
DB 42 SerValaIaaglyThrValgluglyThranglnaspIleSerleuylshphegnluile 61
QY 658 GATCTAAATCAACGACCTGCTCATGAGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 717
DB 62 AspleuthrSerargProalshlsglyglylsthngluuglyleuSerProlysser 81
QY 718 AAACCAATTTGCTACTGATGAGCGGAGATGTCAACATTAACCTTGAGAAAGCTGACTACTA 777
DB 82 LysProphelaIathraspserglalshetProhlslyleuglylshalaaspleuleu 101
QY 778 AAGGCTATTCAGAAACAATTGATCGCTAACGTCACAGTACGACACTACTTGTGAGTC 837
DB 102 LysalaIleaglInglInleuIlelaaenValHisSerAshnshaspTyrrhegluval 121
QY 838 ATTGATTTTGCAGGAGTGAACACTTACTGATCGAAGCGCAAGGTCTACTTGTGTAC 897
DB 122 IleaspPhealaSerAspalaIathrlIethrasparargsnglylshValTyrrhealaasp 141
QY 898 AAAGATGGTGGTGAACCTTGCGCAACCACTGTGCCAAGATTTTGTGTAACCGGACAT 957
DB 142 LysaspglYserValthrleuProthrlnProvalInglInpheluleuSerglYhls 161
QY 958 GTGCGCGTTAGACCATATTAAGAAAAACAATACAAACCAAGCGAAATCTGTTGATGTC 1017
DB 162 ValInglValargProTyrrlyseglulysproIleglInshnglnhlaIylsserValaIaspval 181
QY 1018 GAATTACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGGTCTCAAA 1077
DB 182 GluTyrrThrValglInphethrProleuanProaspAspshpheargProglYleuyls 201
QY 1078 GATCTAAGCTATTGAAAAACATTAAGTATCGGTGACACCATCAACATCTCAAAATTAACA 1137
DB 202 AspThrlyshleuleuylshreulalIleglYaspThrIlethserInglInleu 221
QY 1138 GCTCAAGACAAGACATTTTAAACAAAACCAACCCGAGCTATACGATTATGAAAGCTGAC 1197
DB 222 AlaeglInaIaInserIleleuashnshsrhIsproglYTyrrThrIleTyrrgluashgasp 241
QY 1198 TCTCTAATCGTCACTCATGACAATGACATTTTCCGTAAGATTTCACATGATCAAGAG 1257
DB 242 SerSerIleValThrHIsaspshnshpIlepheargThrIleleuProkethaspInglu 261
QY 1258 TTTACTTACCGGTGTTAAAAATGGGAAACAAGCTTATAGATCAATAAAAATCTGGTCTG 1317

```

```

DB 262 PheThrTyrrHIsValIylshaenarggluglnalaIatyrgluIleashnshlysserglYleu 281
QY 1318 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTAACGCTTAAAAAAGG 1377
DB 282 AsnglugluIleashnshnthraspleuIleSerglulysTyrrValleuylsYegly 301
QY 1378 GAAAAGCCGATATCATCCCTTGTGATCGCAGTCACTTGAAACGTTCCATCAATACGTT 1437
DB 302 GluYspProTyrraspProPheaspargSerHIsleuYshleuethrIleYsrTyrrVal 321
QY 1438 GATGTCGATACCAACGAATTGCTAAAAAGTACAGAGCTCTTAACAGCTACGAAAGCTAAC 1497
DB 322 AspValashnthrasnngluIleuleuylssergluglnleuethrIalasergluashn 341
QY 1498 TTAGACTTCAGACATTTATACGATCTCTCGGTAAGGCTTAATACTCTACACAATCTC 1557
DB 342 LeuaspPheargAspleuTyrraspProIargaspIylshlyshleuTyrrashashnleu 361
QY 1558 GATGCTTTTGGTATTTATGACCTATACCTTAACCGGAAAAAGTAGAGATAATCAAGATGAC 1617
DB 362 AspAlaphneglYllethetaspTyrrThrleuthrIlglyshValgluashpashnshaspasp 381
QY 1618 ACCAACCGTATCATTAACCGTTTATATGGGCAACCGAACCCGAAAGAGAAATGCTAGCTAT 1677
DB 382 ThrashnargIleIethrValTyrrmetglYlyshargProIuglyluashnalaSerTyrr 401
QY 1678 CATTAGCCTATGATATAAGATCGTTATATCCGAAAGAAAGACGAAAGTTTACAGCTACTG 1737
DB 402 HIsleuValatyrraspIylshaspargTyrrThngluInglulargluValTyrrSerTyrrleu 421
QY 1738 CGTTATACAGGGACACCTATACCTGATTAACCTTAACCTTAACGACGAA 1779
DB 422 ArgTyrrThnglyThrProIleProaspashnproashnshpIyls 435

RESULT 15
US-08-488-940-17
; Sequence 17, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

US-08-488-940-17

Alignment Scores:

Pred. No.:	4, 24e-206	Length:	1194
Score:	2120.50	Matches:	415
Percent Similarity:	96.52%	Conservative:	1
Best Local Similarity:	96.29%	Mismatches:	10
Query Match:	67.66%	Indels:	5
		Gaps:	3

US-09-940-235-11 (1-1782) x US-08-488-940-17 (1-1194)

```

QY 496 CAGACCAATCCAGC-----GATCTGGCCCTTCACCGATGTTGTAATGCTGA 546
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 366 GlnThrAsnSerSerValProGlyArgGlySerIle---GluGlyArgGlyLeuIle 384
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 547 CCGATGCTGCTGTATACCGTCATCTGTCAACAAGCCCAATGTTGTTAGGCTGCT 606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 385 ProGluTrpLeuLeuSerPalaProSerValAsnAsnSerGlnLeuValIleValIle 404
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 607 GGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATCGATCTACA 666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 405 GlyThrValGluGlyThrAsnGlnAspIleSerLeuAlaPhePheGlnIleAspLeuThr 424
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 667 TCACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAAGTCCAAATCAAAACATTT 726
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 425 SerAlaProAlaHisGlyGlyAlaThrGluGlnGlyLeuSerProAlaSerIleValProPhe 444
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 727 GGTACTGATAGTGGCCGATGTCACTAACTTGAGAAAGCTGACTTAAAGGCTAAT 786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 445 AlaThrAspSerGlyAlaMetSerHisIleGluGlnIleAspLeuLeuValIle 464
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 787 CAAGAAACAATGATGCTAACGTCACAGTAAGACACTTGAAGTCAATTTGATTTT 846
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 465 GlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlnValIleAspPhe 484
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 847 GCAAGCGATGCAACCTTACTGATCGAAACGCAAGGCTTACTTGTGACAAAGATGAT 906
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 485 AlaSerAspAlaThrIleThrAspArgAsnGlyIleValTyrPheAlaAspIleValAspGly 504
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 907 TCGGTAACTTGGCCGACCCAACTGTCCAAAGATTTTGGTAAAGCGGACATGTGGCGGT 966
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 505 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 524
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 967 AGACCATTAAGAAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTGAATATACT 1026
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 525 Arg---TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1027 GTACAGTTTACTCCCTTAACCTGATGACGATTTTCAGACGAGTCTCAAGATACTAAG 1086
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 544 ValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLysAspThrLys 563
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1087 CTATTGAAGAACTAGCTTACGCTGACACCATCACAATCTCAAGATTTACTAGCTCAAGCA 1146
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeuAlaGlnAla 583
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1147 CAAGCATTTTAAACAAACCAACCAAGGCTATAGATTTATGAAGCTGACTCTCAATC 1206
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 584 GlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArgAspSerSerIle 603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1207 GTCACTCAATGACATGATTTTCCGTAGATTTTACCAATGATGACAGGTTTACTTAC 1266
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 604 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyr 623
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1267 CGTGTAAAGAAATCGGGAAACAAGCTTATAGATCAATAAAATCTGGTGAATGAAGAA 1326
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 624 ArgValLysAsnArgGluGlnAlaTyrArgGlyLeuAsnLysSerGlyLeuAsnGluGln 643
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1327 ATAAACAACACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGGGAAAAAGCCG 1386
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 644 IleAsnAsnThrAspLeuIleSerGlyLysTyrTyrValLeuLysLysGlyGluLysPro 663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1387 TATGATCCCTTTCATCGACGTCACTTGAAACGTTCACATCAAAATACGTTGATGTCAT 1446

```

```

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 664 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 683
QY 1447 ACCAAGCAATGTGCTAAAGAGGACAGCTCTTAACAGCTAGGAAAGCTTAACCTTC 1506
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 684 ThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGluArgAsnLeuAspPhe 703
QY 1507 AGAGATTATAGATCCCTGATAGGCTTAAAGCTTAACTCTCAACAATCTCGATCTTTT 1566
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 704 ArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPhe 723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1567 GGTATTATGACTATACCTTAACTGAAAGATGAGATTAATCAGATGACACCAACCGT 1626
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 724 GlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAspThrAsnArg 743
QY 1627 ATCATACCGTTTATATGAGGCAAGCCGCAAGGAGGAAATGCTAGCTATCATTTAGCC 1686
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 744 IleIleThrValTyrMetGlyLysArgProGluGlnGluAsnAlaSerTyrHisLeuAla 763
QY 1687 TATGATTAAGATCGTTATACCGAAGAAAGAGAGTTTACAGCTACCTGCTTATACA 1746
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 764 TyrAspLysAspArgTyrThrGluGlnGluLysValTyrSerTyrLeuArgTyrThr 783
QY 1747 GGAACACTATACCTGATTAACCTTAACGACAAA 1779
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 784 GlyThrProIleProAspAsnProAsnAspLys 794

```

Search completed: January 28, 2006, 02:41:40
 Job time : 48.4212 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:51:16 ; Search time 116.535 Seconds
 (without alignments)
 12778.538 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134
 Sequence: 1 tcgcctcacgtcgtcgtcg.....atacctaagcagaataa 1782

Scoring table:
 BLOSUM62
 Xgapop 10.0, Xgapext 0.5
 Ygapop 10.0, Ygapext 0.5
 Fgapop 6.0, Fgapext 7.0
 Delop 6.0, Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Command line parameters:
 MODEL=frame_n2p.model -DEV=slp
 O=/cgn2_1/USPTO_bpool_p/US09940235/runat_27012006_144219_27635/app_query.fasta_1.7708
 DB=Published Applications_AA_Main -OPMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0
 -MAXLEN=200000000 -USRR=US09940235 @CGN_1_1.805 @runat_27012006_144219_27635
 -NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NSG_SCORES=0 -WAIT -DSPBLOCK=100
 -LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158	68.9	414	3	US-09-940-235-2
2	2153	68.7	413	4	US-10-360-101-264
3	2132	68.0	414	4	US-10-300-215-252
4	2132	68.0	415	5	US-10-988-943-1
5	2121	67.7	414	4	US-10-300-215-253
6	2090.5	66.7	413	3	US-09-919-703-12
7	1937	61.8	440	5	US-10-474-792-658
8	504.5	16.1	2320	4	US-10-236-392-2
9	504.5	16.1	2328	4	US-10-171-311-64
10	504.5	16.1	2328	4	US-10-236-031B-70
11	504.5	16.1	2328	4	US-10-374-979-98
12	504.5	16.1	2328	4	US-10-182-936A-98

13	504.5	16.1	2328	5	US-10-477-238A-677	Sequence 677, App
14	504.5	16.1	2328	5	US-10-680-287A-677	Sequence 677, App
15	504.5	16.1	2328	5	US-10-477-173-677	Sequence 677, App
16	503	16.0	463	4	US-10-144-194A-52	Sequence 52, Appl
17	503	16.0	463	5	US-10-491-566-52	Sequence 52, Appl
18	503	16.0	642	4	US-10-741-601-354	Sequence 354, App
19	503	16.0	642	5	US-10-741-601-1066	Sequence 1066, App
20	503	16.0	657	4	US-10-741-601-359	Sequence 359, App
21	503	16.0	657	5	US-10-741-601-1072	Sequence 1072, App
22	503	16.0	984	4	US-10-741-601-356	Sequence 356, App
23	503	16.0	984	5	US-10-741-601-1069	Sequence 1069, App
24	503	16.0	2220	4	US-10-236-392-4	Sequence 4, Appl1
25	503	16.0	2296	4	US-10-741-601-363	Sequence 363, Appl
26	503	16.0	2296	5	US-10-741-600-1075	Sequence 1075, App
27	503	16.0	2320	4	US-10-279-733-8	Sequence 8, Appl1
28	503	16.0	2355	4	US-10-144-194A-104	Sequence 104, App
29	503	16.0	2355	4	US-10-360-101-235	Sequence 235, App
30	503	16.0	2355	4	US-10-447-161-3	Sequence 3, Appl1
31	503	16.0	2355	4	US-10-734-564-94	Sequence 94, Appl
32	503	16.0	2355	4	US-10-741-601-357	Sequence 357, App
33	503	16.0	2355	4	US-10-741-601-366	Sequence 366, App
34	503	16.0	2355	5	US-10-491-566-104	Sequence 104, App
35	503	16.0	2355	5	US-10-741-600-1067	Sequence 1067, App
36	503	16.0	2355	5	US-10-741-600-1078	Sequence 1078, App
37	503	16.0	2355	5	US-10-852-335A-147	Sequence 147, App
38	503	16.0	2355	5	US-10-287-436A-436	Sequence 436, App
39	503	16.0	2355	5	US-10-287-436A-1137	Sequence 1137, App
40	503	16.0	2355	6	US-11-040-130-28	Sequence 28, Appl
41	503	16.0	2386	3	US-09-961-403-1	Sequence 1, Appl1
42	503	16.0	2386	4	US-10-741-601-360	Sequence 360, App
43	503	16.0	2386	4	US-10-788-792-206	Sequence 206, App
44	503	16.0	2386	5	US-10-618-281-32	Sequence 32, Appl
45	503	16.0	2386	5	US-10-741-600-1071	Sequence 1071, App

ALIGNMENTS

RESULT 1
 US-09-940-235-2
 ; Sequence 2, Application US/09940235
 ; Publication No. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammarra
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Streptococcus equisimilis
 US-09-940-235-2
 Alignment Scores:
 Pred. No.: 3.4e-192
 Score: 2158.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Length: 414
 Matches: 414
 Conservative: 0
 Mismatches: 0

Query Match:	68.86%	Indels:	0
DB:	3	Gaps:	0
US-09-940-235-11	(1-1782)	x	US-09-940-235-2 (1-414)

QY	538	ATTCTGGAACCTGAGTGCCTGCTGTAACCGTTCATCTGTCACAAACAGCAATTTGGTGT	597
Db	1	ITLeIaegIyProGIuITTrPLeuLeuAspArgProSerValAsnAsnSerGIuLeuValVal	20
QY	598	AGCGTTGCTGTAAGGAGGAGCAATCAAGCATTAAGTCTTTAAATTTTGAATC	657
Db	21	SerValIaGIyThrValGIuGIyThrAsnGlnAspIleSerLeuIyPhePheGIuIle	40
QY	658	GATCTAACATACACACCTGTCATCGAGAAAGACAGACCAAGGCTTAAGTCCAAATCA	717
Db	41	AspLeuThrSerArgProIaHisIeGIyGIyLysThrGIuGlnGlyLeuSerProLysSer	60
QY	718	AAACCAATTTGCTACTGATAGTGGCCGATGTCACATTAACCTTGAGAAAGCTGACTTACTA	777
Db	61	LysProPheIaThrAspSerGIyAlaIaetSerHisIyIleuGIuIyValIaAspLeuLeu	80
QY	778	AAGCTATTTCAAGAACAAATTGATCGCTAAACGTCCACAGTAACGACGACTTGTAGGTC	837
Db	81	LysIaIaIeGIuGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGIuVal	100
QY	838	ATTGATTTTGGAGGAGTGAACCACTTAATCTGATCGAAACGGCAAGCTTACTCTTGCTCAC	897
Db	101	ITLeAspPheIaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp	120
QY	898	AAAGATGGTTCCGTAAACCTTGCCGACCCAAACCTGTCGAAGAATTTTCTAAGCCGACAT	957
Db	121	LysAspGIySerValIThrLeuProThrGlnProValGIuGlnIyPheLeuLeuSerGIyHis	140
QY	958	GTGGCGCTTAAACCATATTAAGAAACCAATCAAAAACCAAGCAATCTGTTGATGTG	1017
Db	141	ValArgValArgProTyrZyIyGIuIySerProIeGIuAsnGlnAlaLysSerValAspVal	160
QY	1018	GAATATATCTGATACAGTTTACTCCCTTAAACCTGATGACATTTGACGACAGGTCGAA	1077
Db	161	GIuTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGIyLeuLys	180
QY	1078	GATACATAAGCTATTTGAAAAACACTAGCTTACGGTGAACACATCAACATCAAGAATTACTA	1137
Db	181	AspThrIySerLeuLysThrIleuAlaIleGIyAspThrIleThrSerGIuLeuLeu	200
QY	1138	GCTCAAGACCAACGATTTTAAACAAAACCAACCGAGCTATACATTTATGAACGTAC	1197
Db	201	AlaGlnIaGlnSerIleLeuAsnLysAsnHisProGIyTyrThrIleTyrGIuArgAsp	220
QY	1198	TCCCTCAATCCGACCTCATGACAAATGACATTTTCCGATGATTTTACCAATGGATCAAGG	1257
Db	221	SerSerIleIleValIThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGIuGln	240
QY	1258	TTTAACCTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATATAAAAAATCTGGCTCG	1317
Db	241	PheThrTyrArgValIyLysAsnArgIuGlnIaIaTyrArgIleAsnLysSerGIyLeu	260
QY	1318	AATGAGAAGAAATTAACAACATGACCTGATCTCTGGAATATTAAGCTTAAAAAAGG	1377
Db	261	AsnGIuGlnIleAsnAsnThrAspLeuIleSerGIyLysTyrValLeuLysIySGly	280
QY	1378	GAAGAAGCGTATGATCCCTTTGATCGACGTCACTTGAACGTTCACCATCAAAATACGT	1437
Db	281	GIuLysProTyrArgAspProPheAspArgSerHisIleuLysIleuPheThrIleLysTyrVal	300
QY	1438	GATGTGCATACCAACGAATTCGCTAAAAAGTAGACAGCTCTTAAACAGCTAGCGACGTAC	1497
Db	301	AspValIAspThrAsnGlnLeuLeuLysSerGIuGlnLeuLeuThrAlaSerGIuArgAsn	320
QY	1498	TTTAACTTCAAGATTTATACGATCCCTGGTGTATAGGCTAAACATCTCTTACAACAATCTC	1557
Db	321	LeuAspPheAspLeuTyrArgProArgAspLysValAlaLysIleuLysTyrAsnAsnLeu	340

QY 1558 AAATGCTTTTGGATATATGACATATACCTTAACCTGAAAAGTAGAGATTAATTCAGATGAC 1617
 Db 341 ASpAaApNeGlyLTaMeCaSPryrThrLeuThGlyLVaValGluSpaSnH:saSPaSP 360
 QY 1618 ACCAACCGTATCATCAACGCTTATATATGGCGAAGCGACCCGAGAGAGAAATGCTAGCTAT 1677
 Db 361 ThraMaAgiLeIeThrValTyrMeTgLySaRgProGluGlyGluMaNaLaSeTyr 380
 QY 1678 CATTAGCCCTATGATTAAGAATGTTATACCGAAGAAAGACGAGAGCTTTACAGCTACTG 1737
 Db 381 HisLeuAlaTyrASpLysaSPaRgTyrThrGluGluValaGlyValTyrSeTyrLeu 400
 QY 1738 CGTTATACAGGAGACACTATACCTGATTAACCTTAACGACAA 1779
 Db 401 ArgTyrThrGlyThrProLeProAspaSnProAspaLys 414
 RESULT 2
 US-10-360-101-264
 ; Sequence 264, Application US/10360101
 ; Publication No. US20040009550A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moll, Gert N.
 ; APPLICANT: leenhouts, Cornelis J.
 ; TITLE OF INVENTION: Export and modification of (poly) peptide in the lantibiotic way
 ; FILE REFERENCE: 2183-5673
 ; CURRENT APPLICATION NUMBER: US/10/360,101
 ; CURRENT FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: EP 02077060.8
 ; PRIOR FILING DATE: 2002-05-24
 ; NUMBER OF SEQ ID NOS: 309
 ; SOFTWARE: PatentIn version 3.1.
 ; SEQ ID NO 264
 ; LENGTH: 413
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: sequence of streptokinase
 US-10-360-101-264
 Alignment Scores:
 Pred. No.: 9.97e-192 Length: 413
 Score: 2153.00 Matches: 413
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.70% Indels: 0
 DB: 4 Gaps: 0
 US-09-940-235-11 (1-1782) x US-10-360-101-264 (1-413)
 QY 538 ATGTCGTGACCTGAGTGGCTGTCTAGACCGTCCATCGTCAACAACAGCCATTGTTGTT 597
 Db 1 lIeaIaGlyProGluThrLeuLeuAspaRgProSeValaSnAsnSeRgInLeuVala 20
 QY 598 AGCGTTGCTGACTCTGTTGAGGGACGAATCAAGACATTAGTCTTAATTTTGAATC 657
 Db 21 SeValaIaGlyThrValaGluGlyThraSngInaSPilSeSeRleuLyspHeGluIle 40
 QY 658 GATCAATATACACGACCTGCTCATGAGGAAGAAGACGAGAAGCTTAATCCAAATCA 717
 Db 41 AspLeuThrSeArpProAlaHISgLyGlyLysThGluGluInGlyLeuSerProLysSeR 60
 QY 718 AAACCATTTGCTACGATATAGTGGCGGATGTCAATTAACCTTAGAAGCTGACTTACTA 777
 Db 61 LysProPheAlaThrAspSerGlyAlaMeTSeHisLysLeuGluLysAlaAspLeu 80
 QY 778 AAGGCTATTCAAGAACATATGATCGCTAAAGTCCACAGTAAACGACGACTATTGAGGTC 837
 Db 81 LysAlaIaIeGInGluGluInLeuIleAlaAsnValaHISeTasAspaSPryrPheGluVala 100
 QY 838 ATTGATTTTGGAAAGGAGGCAACCATTTACTGTGATTCGAAACGGCAAGGCTACTTTGGCTAC 897
 Db 101 lIeaSPpHeaLaSeTAspaLaThrIleThrAspaRgAsnGlyLysValTyrPheAlaSP 120

```

QY 898 AAAAGATGCTTCGGTAACTTCCGACCAACCTGTCACAAATTTTGTAAAGCGGACAT 957
DB 121 LysAspGlySerValThrLeuProthGlnProValGlnGlnPheLeuSerGlyHis 140
QY 958 GTGCGGGTTAGACCATTAAGAAAAACAATACAAACCAAGGAAATCTGTGATGTC 1017
DB 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATACGTATACAGTTTACTCCCTTAAACCCGTATGACGATTTCAAGCAGTCTCAAA 1077
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTTTAAACGTCAC 1197
DB 201 AlaGlnAlaIleGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 220
QY 1198 TCTCATGCTGTCATCATGACATGACATTTTCCGTACGATTTTACCAAGATCAAGAG 1257
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1258 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCGTCTG 1317
DB 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTTGAGAAATATTAAGTCTTAAAAAGG 1377
DB 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValIleLysLysGly 280
QY 1378 GAAAGACCGTATGATCCCTTGTATGCGAGTCACTTGAAATGTCACCAATAAGT 1437
DB 281 GluLysProTyrAspProPheAspArgSerHisLysLeuPheThrIleLysTyrVal 300
QY 1438 GATGTCGATACCAACAATTGCTAAAGAGACAGCTCTTAACAGCTACGAAACGTAAC 1497
DB 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysAsn 320
QY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTTAACCTACTTCAACAATCTC 1557
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 340
QY 1558 GATGCTTTGGTATTAATGACATTAACCTTAACGAAAAAGTAGAGATATACAGATGAC 1617
DB 341 AspAlaPheGlyLysLeuAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1618 ACCAACCGTATCACTAACCGTTATATGAGGACGACCCGAGAGAGAGAAATGCTAGCTAT 1677
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAlaSerTyr 380
QY 1678 CATTTAGCCTATGATTAAGATGCTTATACGAGAAAGAAAGAAAGATTTACAGTACTG 1737
DB 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnLysValTyrSerTyrLeu 400
QY 1738 CGTTATACGGAACACTTATCTGATTAACCTTAACGAC 1776
DB 401 ArgTyrThrGlyThrProIleProAspAsnProAspAsn 413

```

RESULT 3

```

US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215

```

```

; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252
Alignment Scores:
Pred. No.: 9,12e-190 Length: 414
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 98.03% Indels: 0
DB: Gaps: 0
US-09-940-235-11 (1-1782) x US-10-300-215-252 (1-414)
QY 538 ATTGCTGACCTGATGCTGCTGACACCGTCCATCTGCAACAACGCAATGATGTT 597
DB 1 IleAlaGlyProGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 598 AGCGTTCGTGATCTGTTGGGGGACGAATCAACATTAAGCTTAAATTTTGAATC 657
DB 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPheGlnLys 40
QY 658 GATTAACATCAACGACCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 717
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnLysLeuSerProLysSer 60
QY 718 AAACATTTGCTACTGATAGTGGCGGATGTCACATTAACCTTGAAGAACTGACTACTA 777
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 80
QY 778 AAGCTATTCAGAAACAATGATGCTTAAGCTCAACAGTAAACGACTTTGAGGTC 837
DB 81 LysAlaIleGlnGlnGlnLysLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100
QY 838 ATTGATTTTGAACGATGACCAACATTAAGTGAAGAAAGGCAAGTCTACTTGTGAC 897
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 898 AAAGATGTTCCGTTAACTTCCGACCAACCTGTCACAAAGAAATTTTGTAAAGCGGACAT 957
DB 121 LysAspGlySerValThrLeuProthGlnProValGlnGlnPheLeuSerGlyHis 140
QY 958 GTGCGGGTTAGACCATTAAGAAAAACAATACAAACCAAGGAAATCTGTGATGTC 1017
DB 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1018 GAATATACGTATACAGTTTACTCCCTTAAACCCGTATGACGATTTTCAAGCAGTCTCAAA 1077
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1137
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200

```

QY 1138 GCTCAAGCAAAAGCATTTTAAACAAAACCAAGCGCTATACGATTATGAACTGAC 1197
Db 201 AAGlnAlaGlnserlIeLeuAsnlySThrHISProGlyTyrThrIleTyrGluArgAsp 220
QY 1198 TCCTCAATCGTACATGACAAATGACATTTTCCGATGATTTTACCAATGATGACAG 1257
Db 221 SerSerIleValIthrhISAspAsnspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAATCGTCTG 1317
Db 241 PheThrTyrHISValIlyAsnAsnArgGlnGlnAlaTyrGluIleAsnlySlySerGlyLeu 260
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTGTAGAAAATATTATTCGCTTAAAAAAAGG 1377
Db 261 AsnGlnGluIleAsnAsnThrAspLeuIleSerGlnIlyTyrIleValIleuIlySlyGly 280
QY 1378 GAAAGCCGATGATCCCTTGTATGCGAGTCACTTGAATCTGTACCAATACGTT 1437
Db 281 GlnIlySProTyrAspProPheAspArgSerHISleuIlySleuPheThrIleIlySlyVal 300
QY 1438 GATGTCGATACCAAGCAATTCGTAATAAAGTGAAGCAGCTTAAAGCTGAGCAAGTAA 1497
Db 301 AspValAsnThrAsnGlnLeuLeuIlySerGlnGlnLeuLeuThrAlaSerGlnArgAsn 320
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGCTTAACTATCAACAATCTC 1557
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspIlySAlaIlySleuLeuTyrAsnAsnLeu 340
QY 1558 GATGCTTTTGGTATTATGACATTAACCTTAACTGAAAAGTGAAGATTAATCAAGTAC 1617
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyIlySValGlnAspAsnHISAspAsp 360
QY 1618 ACCAACCGTATCATACCGTTATATATGCGCAAGCAAGCGAAGAGAAAGTACTACT 1677
Db 361 ThrAsnArgIleIleThrValIlyMetGlyIlySArgProGlnGlyGlnAsnAlaSerTyr 380
QY 1678 CATTTAGCCTATGATTAAGATCGTTATACCAAGAAAGAGAAAGTTTACAGCTACG 1737
Db 381 HisLeuAlaTyrAspIlyAspArgTyrThrGlnGlnGlnIlyValIlySerTyrLeu 400
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAA 1779
Db 401 ArgTyrThrGlyTyrThrProIleProAspAsnProAsnAspIlyS 414

RESULT 4

US-10-988-943-1
; Sequence 1, Application US/10988943
; Publication No. US20050176085A1
; GENERAL INFORMATION:
; APPLICANT: Center for Genetic Engineering and Biotechnology
; TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
; TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
; FILE REFERENCE: Proteomics CU2003-269
; CURRENT APPLICATION NUMBER: US/10/988,943
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-988-943-1

Alignment Scores:

Pred. No.: 9,136-190 Length: 415
Score: 2132.00 Matches: 409
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 4
Query Match: 68.03% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-11 (1-1782) x US-10-988-943-1 (1-415)

QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTTCATCTGTCAACAACAGCAATGGTGT 597
Db 2 IleAlaGlyProGlnIlyProIleuLeuAspArgProSerValAsnAsnSerGlnLeuVal 21
QY 598 AGCGTGTCTGTACCTGTGAGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 657
Db 22 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuIlySlyPheGluIle 41
QY 658 GATCTAACATCAGACCTGTGTATGAGGAAACAGAGAAAGCTTAAAGTCCAAATCA 717
Db 42 AspLeuThrSerArgProAlaHISGlyIlySThrGlnGlnIlyLeuSerProIlySer 61
QY 718 AAACATTTTCTAGTAGTGGCGCATGTCACTTAACTTGAAGAGCTTACTA 777
Db 62 LysProPheAlaThrAspSerGlyAlaMetProHISlySleuGlnIlySAlaAspLeu 81
QY 778 AAGGCTAATTGAAGCAATTTGATTCGCTAACGTCCACAGTACGACTATTGAGGT 837
Db 82 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHISerAsnAspArgTyrPheGluVal 101
QY 838 ATTGATTTTGCAGCGATGCAACATTTATCTGATCGAAACGGCAAGCTTACTTGTGAC 897
Db 102 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlySValTyrPheAlaAsp 121
QY 898 AAGAGTGTTCGGTAACTTCGCGACCCGACCTGTCCAGAAATTTTGTAAAGCGACAT 957
Db 122 LysAspGlySerValThrLeuProHISProIlyGlnIlyPheLeuSerGlyHIS 141
QY 958 GTGCGGTTAGACCATTAATAAAGAAAACAATACAAAACCAAGCAAAATCTGTGATGT 1017
Db 142 ValArgValArgProTyrIlySglnIlySProIleGlnAsnGlnAlaIlySserValAspVal 161
QY 1018 GAATATACCTGTACAGTTTACTCCCTTAAACCTGTATGACATTTACAGCAAGCTCTCAA 1077
Db 162 GlnTyrThrValGlnPheThrProLeuAspProAspAspPheArgProGlyLeuIlyS 181
QY 1078 GATACTAAGCTATTTGAAAACATCAGTATGCGGACCACTCATCTCAAGTAAATTA 1137
Db 182 AspThrLysLeuLeuIlySThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 201
QY 1138 GCTCAAGCAAAAGCAATTTTAAACAAAACCAAGCGCTATGACATTTTGAACGTGAC 1197
Db 202 AlaGlnAlaGlnserIleLeuAsnlySThrHISProGlyTyrThrIleTyrGlnArgAsp 221
QY 1198 TCCTCAATCGTCACTGATGCAATGACATTTTCCGATGATTTTACCAATGATCAAG 1257
Db 222 SerSerIleValIthrhISAspAsnspIlePheArgThrIleLeuProMetAspGlnGlu 241
QY 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGGATCAATAAAAATCGTCTG 1317
Db 242 PheThrTyrHISValIlyAsnAsnArgGlnGlnAlaTyrGlnIleAsnlySlySerglyLeu 261
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTGTAGAAAATATTATTCGCTTAAAAAAAGG 1377
Db 262 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnIlySlyTyrIleValIleuIlySlyGly 281
QY 1378 GAAAGCCGATGATCCCTTGTATGCGAGTCACTTGAATCTGTACCAATACGTT 1437
Db 282 GlnIlySProTyrAspProPheAspArgSerHISleuIlySleuPheThrIleIlySlyVal 301
QY 1438 GATGTCGATACCAAGCAATTTGTAATAAAGTGAAGCAGCTTAACTGATGACGAACTGAC 1497
Db 302 AspValAsnThrAsnGlnLeuLeuIlySerGlnGlnLeuLeuThrAlaSerGlnArgAsn 321
QY 1498 TTAGACTTCAGAGATTTATACGATCCTCGTGAATAAGCTTAACTATCAACAATCTC 1557
Db 322 LeuAspPheArgAspLeuTyrAspProArgAspIlySAlaIlySleuLeuTyrAsnAsnLeu 341
QY 1558 GATGCTTTTGGTATTATGACATTAACCTTAACTGAAAAGTGAAGATTAATCAAGTAC 1617
Db 342 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyIlySValGlnAspAsnHISAspAsp 361

QY 1618 ACCAACCGTATCATACCGTTTATATGGCGACGCCGAGAGAGAAATGCTAGCTAT 1677
Db 362 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluValAsnAlaSerTyr 381
QY 1678 CATTTAGCCTATGATTAAGAATGCTTATACCGAAGAAAGAGAGATTACCTACTG 1737
Db 382 HisLeuAlaIleTyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 401
QY 1738 CGTTATACAGGAGACACCTATACCTGATTAACCTTAACGACAA 1779
Db 402 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 415
RESULT 5
US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-CON.1
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified strep protein
US-10-300-215-253.
Alignment Scores:
Pred. No.: 9,71e-189 Length: 414
Score: 2121.00 Matches: 407
Percent Similarity: 98.79% Conservative: 2
Best Local Similarity: 98.31% Mismatches: 5
Query Match: 67.68% Indels: 0
Gaps: 0
US-09-940-235-11 (1-1782) x US-10-300-215-253 (1-414)
QY 538 ATTGCTGACCTGAGTGGCTGTAGACCTGTCATCAACAAGCCAAATGGTTGTT 597
Db 1 IleAlaGlyProGluIleTyrLeuAspArgProSerValAsnAsnSerGluLeuVal 20
QY 598 AGCGTTGCTGATCTGTGAGGGGAGCAATCAAGATTAAGTCTTAATTTTGAATC 657
Db 21 SerValAlaGlyThrValGluGluThrAsnGluAspIleSerLeuLysPhePheGluIle 40
QY 658 GATTAACATCAAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAAGTCCAAATCA 717
Db 41 AspLeuThrSerArgProIleHisGlyGlyLysThrGluGluIleLysLeuSerProLysSer 60
QY 718 AAACCATTTGCTACTGATAGTGGCCGCGATGTCACATTAACCTTGAGAAAGCTGACTTA 777

Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluValAspLeuLeu 80
QY 778 AAGGCTATTCAAGAAACAATTGATCGCTTAACCTCCACATTAACGACATTAAGTTC 837
Db 81 LysAlaLysGluGluGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 838 ATTGATTTTTCGAAACGATGCAACCATTAAGTATGCAAAACGGCAAGCTACTTGGCTAC 897
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 898 AAAGATGCTTCGTTACCTTGCCGACCCGACCTGTCGAAAGATTTTTCGACCGACAT 957
Db 121 LysAspGlySerValThrLeuProThrGluProValGluGluPheLeuLeuSerGlyHis 140
QY 958 GTGGCGCTTAGACCATTAATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1017
Db 141 ValArgValArgProTyrLysGluLysProIleGluAsnGluAlaLysSerValAspVal 160
QY 1018 GAATATAGCTGATGAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGCTCAAA 1077
Db 161 GluTyrThrValGluPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1078 GATATCAAGCTATTGAAAACACTGACTTCGTGACACCATCACATCTCAAGAAATTACTA 1137
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGluGluLeuLeu 200
QY 1138 GCTCAAGACAAAGACATTTTAAACAAAACCAACCGGCTTACGATTTTAAACGTGAC 1197
Db 201 AlaGluAlaGluSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 220
QY 1198 TCCGTAATCGCATCTGACCAATGACATTTTCGTCGATTTTACCAATGATCAAGAG 1257
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGluIle 240
QY 1258 TTTACTTACCGTGTGTTAAAAATCGGAAACAAGCTTATGATCAATTAATAAATCGTGTG 1317
Db 241 PheThrTyrHisValLysValHisAsnArgGluGluAlaTyrGluIleAsnLysLysSerGlyLeu 260
QY 1318 AATGAGAAATTAACAACACTGACCTGATCTTGAGAAATTTTACGTCCTTAAAAAGG 1377
Db 261 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
QY 1378 GAAAAGCGGTATGATCCCTTTGATCGGACGATCACTTGAACCTGTTACCATTAATAGTT 1437
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysPheVal 300
QY 1438 GATGCGATACCAACGAATTGCTTAAAAAGTAGAGAGCTTAAACGTAAGGAAAGCTAAC 1497
Db 301 AspValAsnThrHisAsnGluLeuLeuLysSerGluGluLeuLeuThrAlaSerGluArgAsn 320
QY 1498 TTAGACTTCAGAGATTATTAAGATCTTCGTGATAGGCTTAACCTTACCAACATCTC 1557
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1558 GATGCTTTTGCATTAATGACATTAACCTTAACTGAAAAGTAGAGATTAATCACGATGAC 1617
Db 341 AspAlaPheGlyIleMetAspArgTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1618 ACCAACCGTATCATACCGTTTATATGGGACGCCGAGAGAGAAATGCTAGCTAT 1677
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluValAsnAlaSerTyr 380
QY 1678 CATTTAGCCTATGATTAAGAATGCTTATACCGAAGAAAGAGAGATTACCTACTG 1737
Db 381 HisLeuAlaIleTyrAspLysAspArgTyrThrGluGluGluValTyrSerTyrLeu 400
QY 1738 CGTTATACAGGAGACACCTATACCTGATTAACCTTAACGACAA 1779
Db 401 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 414
RESULT 6
US-09-919-703-12


```

; Sequence 12, Application US/09919703
; Patent No. US20020165129A1
; GENERAL INFORMATION:
; APPLICANT: Kravetz, Gerald
; APPLICANT: Rabin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; FILE REFERENCE: 50216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 413
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-12

Alignment Scores:
Pred. No.: 6,84e-186 Length: 413
Score: 2090.50 Matches: 408
Percent Similarity: 98.80% Conservative: 2
Best Local Similarity: 98.31% Mismatches: 2
Query Match: 66.70% Indels: 3
DB: 3 Gaps: 3

US-09-940-235-11 (1-1782) x US-09-919-703-12 (1-413)
QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAGCCAAATTGTTGTT 597
DB 1 Ileaaglyprrglttrleuenuaspargproservalasmsnserglneuvalval 20
QY 598 AGCGTGTCTGACTGTGTGAGGGAAGATCAAGCATTAAGTCTTAATTTTGAATC 657
DB 21 SerValaIaglyThrVal---GlyThrANGlnAspIleSerIeuylshphegluile 39
QY 658 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGAGGCTTAAGTCCAAATCA 717
DB 40 AspIeuThrSerArgProIahIasgllylysthglngllyleuSerProlySer 59
QY 718 AAACCATTTGCTACTGATAGTGCGGAGATGTCACATTAACCTTGAGAAAGCTGACTACTA 777
DB 60 LysProPhealathraspserglYalAetSerHlslylsleuglulysAlaaspleu 79
QY 778 AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTACGACGACTTCTTGAGGTC 837
DB 80 LysAlaIleIegInglnglneuleleahsnValHisSerAsnhsaspslyrPhegluVal 99
QY 838 ATTGATTTTGAAGGAGTCAACATTAAGTCAAGACGCGAAGGCTTACTTGTCTGAC 897
DB 100 IleasPhealaserAspAlaThrIleThrAspArgAsnglylysvAllyrPhealAasp 119
QY 898 AAAGATGTTGGGTAACTTGCGGACCCCAACCTGTCGCAAGATTTTGTGTAAGCGGACAT 957
DB 120 LysaspGlySerValThrIeuProThrGlnProValGlnGlnPheIeuSerGlyHis 139
QY 958 GTGCGCGTTAAGACATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTC 1017
DB 140 ValaIagValaIargProIylrlysglylsrProlIegInasnglnhAlaIyaserValaIaspVal 159
QY 1018 GAATATCTGTACAGATTACTCCCTTAAACCTGATGACGATTTCAGACAGGCTCTCAAA 1077
DB 160 GluTyThrValaIglhPheThrProIeuaenProAspAspAspAspAspAspAspAspAsp 179
QY 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACATCATCATCTCAAGATTACTA 1137

```

```

DB 180 LeuThrIlyseuIeuylsrThrIleAlaIleGlyAspThrIleThrSerGlnIleuIeu 199
QY 1138 GCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGCGTATACGATTATGAAACGTCAC 1197
DB 200 AlaIglhAlaIasrIleleuAsnlysaenhsIeProIylThrIleIyrgIuArgAsp 219
QY 1198 TCCCATGCTGCACTACATGACAAATGACATTTCCGTACAGATTATTAACCAAGATCAAGAG 1257
DB 220 SerSerIleValaThrIhsaPasnAspIlePheArgThrIleIeuProIeuaenhsIe 239
QY 1258 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATGATCAATATAAAAAATCGTCTG 1317
DB 240 PheThrTyArgValaIylsaenhsIeuglnAlaIyArgIleAsnlyslsSerGlyIeu 259
QY 1318 AATGAAGAAATTAACAACATGACCTGATCTTGAGAAATAT--TACGCTCTTAAAAA 1374
DB 260 AsnGlnGlnIleAsnAsnThrAspIleuIleSerIeuIuTyArgIyValaIleuyls 279
QY 1375 GGGGAAAAAGCCGATGATCCCTTGATCGGACGCTTGAACCTGTTACACATCAATAC 1434
DB 280 GlyIuIyepProIylrAspProPheAspArgSerHlsIeuIySeuPheThrIleIyArg 299
QY 1435 GTTGATGCTGATACCAACGAATTTGCTAAAAGTGAAGAGCTTTAACAAGCTAGCGAAGCT 1494
DB 300 ValaIspValaIasPthrANGlnIleuIeuIySerGlnIleuIeuThrAlaSerGluArg 319
QY 1495 AACTTACGCTTACAGATTTATACATCTCTGTGATACGCTTAACTTACATCTTACACAT 1554
DB 320 AsnIleuAspPheArgAspIeuTyArgProIArgAspIyAlaIyIleuTyArgAsn 339
QY 1555 CTGAGTCTTTGGTATTTATGACATACTTAACTGGAAGAAAGTAAAGATACAGAT 1614
DB 340 LeuAspAlaPheIleIyIleuAspTyThrIleuThrIlyIyValaIAspAsnhsIasp 359
QY 1615 GACACCAACCGTATCAATTAACGTTTATATGGGCAAGGCAACCGAAGGAGAAATGTAC 1674
DB 360 AspThrAsnhsIleIleThrValIyrmecGlyIyArgProIuglnIyIuAsnhsIser 379
QY 1675 TATCATTTAGCTTATGATTAAGATCGTTATACCGAAGAACAGAGATTTCAGCTAC 1734
DB 380 TyrHis---AlaTyAspIyAspArgTyThrGlnIuglnIyIuArgIuValaIySerTy 398
QY 1735 CTGCGTTATACAGGAGACCTTATACCTGATTAACCTTACAGCAAA 1779
DB 399 LeuArgTyThrGlnIyThrProIleProAspAsnProAspAspIys 413

RESULT 7
US-10-474-792-658
; Sequence 658, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Louie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; PRIOR FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 658
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-658

Alignment Scores:
Pred. No.: 1.51e-171 Length: 440
Score: 1937.00 Matches: 367
Percent Similarity: 93.48% Conservative: 20
Best Local Similarity: 88.65% Mismatches: 27

```

Query Match:

DB:	5	Gaps:	0
-----	---	-------	---

US-09-940-235-11 (1-1782) X US-10-474-792-658 (1-440)

QY	538	ATTGTGGACCTGAGTGGCTGTCTAAACCGTTCATGCTCAACAACAGCAATTTGGTGT	597
Db	27	ITlelaaglyTyrGlyTripleuproraaprrgProProlleasnhserserGlnleuVal	46
QY	598	AGCGTGTCTGCTACTGTGTGAGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC	657
Db	47	SerleatlaaglyIleValGlnGlyThrAspIysValPheIleasnheheGlnIle	66
QY	658	GATCTAACATCAACGACCTGTCAATGAGAAAGACAGAGCAAGCGTTAAGTCCAAATCA	717
Db	67	AspleuthrSerGlnProIahIselyGlyYsrInGlnGlyLeuSerProIysSer	86
QY	718	AAACCATTTTGGTACTGATAGTGCCGATGTTCACATTAACCTTGAAGAAAGCTACCTACA	777
Db	87	LysProPhealatrAspAsnGlyIalMetProHIslySbleGlnIlySalaaspIeuleu	106
QY	778	AAGCTATTCAAGAACATTTGATCGCTTAACTGCCACAGTAAACGACGACTCTTGAAGCTC	837
Db	107	LysIalIleGlnIlySbleuIleIalasnValIHisSerAsnAspGlyTyrPheGlnVal	126
QY	838	ATTGATTTTGGAAAGGATGCAACCTTAAGTGTGTGAAACGGGAAGGTCTACTTGTCTGAC	897
Db	127	IleasprhealaserAspIalatrIlethrAspAsnGlyIysValTyrPheIlaAsp	146
QY	898	AAABATGGTTCGGTAACCTTGCCGACCCGACCTGTCCAAAGATTTTGTCTAAGCCGACAT	957
Db	147	LysAspGlySerValThrLeuProThrGlnProValGlnIlyPheIleuIlySelyIHis	166
QY	958	GTGGCGGTTAACCATATTAAGAAANACCATATCAAAAACMAAGCAATCTGTGATGTG	1017
Db	167	ValArgValaIspProTyrIlySbleIlySbleProValGlnasnglnIalysSerValaAspVal	186
QY	1018	GAATATACTGTACAGTTACTCCCTTAAACCTGATGCGATTTGACACCGGTCCAA	1077
Db	187	GlnTyrThrValGlnPheThrProleuasnProAspAspAspPheValProGlyIleuIys	206
QY	1078	GATCTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCAACATCTCAAGAAATTACTA	1137
Db	207	AspThrIlySbleuIlySbleThrleuIalIleGlyAspThrIlethrSerGlnIlyleuIeu	226
QY	1138	GCTCAACGACAAACGATTTTAAACAAAAACACCCAGGCTATGACTTATTAAGACGTGAC	1197
Db	227	AlaGlnIalGlnSerIleIeuasnIlySthrIspProGlyTyrThrIleTyrGlnArgAsp	246
QY	1198	TCCCTCAATCGCAGCTGACAAATGACATTTTCCGTAGACTTTTACCAATGATACAG	1257
Db	247	SerSerIleValIThrIleAspAsnAspIlePheArgThrIleleuProMetAspGlnIu	266
QY	1258	TTTACTTAACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGGTCTG	1317
Db	267	PheThrTyrIHisValIlySAsnArgIuGlnIalTyrGlnIleasnProIySthrGlyIle	286
QY	1318	AATGCAAAATTAACAACAGCTGATCTCTGAGAAATTAACGTCTTAAAAAAGG	1377
Db	287	LysGlnIlySthrAsnAsnThrAspIleuValSerGlnIlyTyrValIleuIlySbleGlnIly	306
QY	1378	GAAGAAGCGTATGATCCCTTGAATCGCAGTCACTTGAACGTTCCACATCAAAATACGTT	1437
Db	307	GlnIlySerProTyrAspProPheAspAsnSerIleIeuIlySbleuPheThrIleIySerTyrVal	326
QY	1438	GATGTGCATACCAACGAATTGCTTAAAAAGTAGAGACGCTCTTAACAGCTAAGCAACGTAC	1497
Db	327	AspValAsnThrAsnGlnIleuIlySserGlnIleuIeuThrAlaSerGlnIlyArgAsn	346
QY	1498	TTTGAAGTTCAGAGATTTATAGACCTCGTGATTAAGGTTAAACATCTTACAAACAATCTC	1557
Db	347	LeuAspPheAspIleuTyrAspProArgAspIlySalaIlySbleuIlyTyrAsnAsnIeu	366

Oy	1558	GATCCTTTGGTATTATTAAGACTATACCTTAACGTAAAGAGAGGATTAATCAACGATAC	1617
Db	367	AspAlaIheAspIleMetAspTyrThrIeuThrIyIyValGIuAspAsnHLSAspLys	386
Oy	1618	ACCAACCGTATCATTAACCGTTTATATGAGCAAGCAACCCGAAGAGAGAATGCTAGCTAT	1677
Db	387	AsnAsnArgValValThrValThrValTyrMetGlyIyAsnArgProIySgIlyAlaIySgIySerTyr	406
Oy	1678	CATTAGCCCTTATGATTAAGATCGTTATACGAGAGAAGACGAGATTACAGCTACCTG	1737
Db	407	HisIeuAlaTyrAspLysAspLeuTyrThrGIuGIuIuIuArgLysAlaTyrSerTyrLeu	426
Oy	1738	CGTATACAGGACCACTTATACCTGATTAACCCGTAAAGACAAA	1779
Db	427	ArgAspThrIuGIyThrProIleProAspAsnProLysAspLys	440
RESULT 8			
US-10-236-392-2			
Sequence 2, Application US/10236392			
Publication No. US20040067490A1			
GENERAL INFORMATION:			
APPLICANT: Anderson, David W			
APPLICANT: Boldog, Ferenc L			
APPLICANT: Burgess, Catherine, E			
APPLICANT: Caeman, Stacie J			
APPLICANT: Catterton, Elina			
APPLICANT: Chapoval, Andrei			
APPLICANT: Crabtree, Julie			
APPLICANT: Edinger, Shlomit, R			
APPLICANT: Eilerman, Karen			
APPLICANT: Gerlach, Valerie			
APPLICANT: Gorman, Linda			
APPLICANT: Grose, William M			
APPLICANT: Gusev, Vladimir			
APPLICANT: Kekuda, Ramesh			
APPLICANT: LaRochele, William J			
APPLICANT: Li, Li			
APPLICANT: MacDougall, John R			
APPLICANT: Malyankar, Uriel M			
APPLICANT: Miller, Charles E			
APPLICANT: Miller, Isabelle			
APPLICANT: Padigaru, Muralidhara			
APPLICANT: Patnrajan, Meera			
APPLICANT: Pena, Carol A			
APPLICANT: Peyman, John A			
APPLICANT: Rastelli, Luca			
APPLICANT: Reiger, Daniel K			
APPLICANT: Rothenberg, Mark E			
APPLICANT: Shenoy, Suresh			
APPLICANT: Shinkete, Richard A			
APPLICANT: Smithson, Glenda			
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME			
FILE REFERENCE: 21402-442A			
CURRENT APPLICATION NUMBER: US/10/236, 392			
CURRENT FILING DATE: 2002-09-06			
PRIOR APPLICATION NUMBER: US09/540,763			
PRIOR FILING DATE: 2000-03-30			
PRIOR APPLICATION NUMBER: US60/390,155			
PRIOR FILING DATE: 2002-06-19			
PRIOR APPLICATION NUMBER: US09/635,949			
PRIOR FILING DATE: 2000-08-10			
PRIOR APPLICATION NUMBER: US60/318,765			
PRIOR FILING DATE: 2001-09-12			
PRIOR APPLICATION NUMBER: US60/357,303			
PRIOR FILING DATE: 2002-02-15			
PRIOR APPLICATION NUMBER: US60/367,753			
PRIOR FILING DATE: 2002-03-25			
PRIOR APPLICATION NUMBER: US60/369,479			
PRIOR FILING DATE: 2002-04-02			
PRIOR APPLICATION NUMBER: US09/659,634			
PRIOR FILING DATE: 2000-09-12			
PRIOR APPLICATION NUMBER: US60/318,120			
PRIOR FILING DATE: 2001-09-07			

```
/ PRIOR APPLICATION NUMBER: US60/318,130
/ PRIOR FILING DATE: 2001-09-07
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 794
/ SOFTWARE: Custom
/ SEQ ID NO 2
/ LENGTH: 2320
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-236-392-2

Alignment Scores:
Pred. No.: 1,66e-37 Length: 2320
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: Gaps: 4

US-09-940-235-11 (1-1782) x US-10-236-392-2 (1-2320)

QY 254 CCCATAGCTGAGAGTGTTCATCATGCTGCTGGAGCTTCTATGTGTCGAGAAACG 313
Db 146 Protlealaglulycyspheasphsalaalaglythrseryvalvalglulthr
QY 314 TGGGA-----GAAGCAGC 327
Db 166 TrrglulysprotyrclnglytrpmetmetvalasrcysthrCysleuglyglulysr 185
QY 328 GAGCGCATCACTTGACCTTCTAGAAATAGATGCAACATCAGACACAGACATCTAT 387
Db 186 GlyargilethrCysThrserArgasnArgCysasnserglinspThrargThrserTy 205
QY 388 AGAATTGAGACACCTGAGCAAGAGATTAATCGAGAAACCTGCTCGATGCAATGTC 447
Db 206 ArgileglyasprThrtrpserlylsaspsasnArglyAsnleuenglncysileCys 225
QY 448 ACAAGCAACGCGCGAGAGAGAGTGAGAGTGTGAGAGGACACCTCTGTGCAACCATG 507
Db 226 ThrGlyasnGlyArgGlyGlyutrrplysCysgluArgHsThrserValGlnThrThrser 245
QY 508 AGCGGATCTGGCCCTTCAACCGATGTCGATTGCT----- 543
Db 246 SerGlyserGlyProphethrAspValArgAlaValTyrGlnProGlnProHsPro 265
QY 544 -----GACCT 549
Db 266 GlnProProtyrGlyHisCysValThrAspserGlyValValTyrserValGlyMet 285
QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACACAGCCAAATGGTTGTTAGCTGCTGT 609
Db 286 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysleugly 302
QY 610 ACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATTCGATCAACATCA 669
Db 303 AsnGlyValserCysGlnGlnThrAlaValThrGlnThrTyrGlyAsnleuasnGly 322
QY 670 GACCT-----GCTCATGAGAGAAAGCA 693
Db 323 GluProCysValleuProphethrTyrAsnGlyArgThr 335

RESULT 9
US-10-171-311-64
/ Sequence 64, Application US/10171311
/ Publication No. US20030087270A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Chen, Yan
/ APPLICANT: Zhao, Xumei
/ APPLICANT: Monahan, John
/ APPLICANT: Kamackar, Shubhangi
/ APPLICANT: Glatc, Karen
/ APPLICANT: Gannavarapu, Manjula
```

```
/ APPLICANT: Hoersth, Sebastian
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
/ TITLE OF INVENTION: OF CERVICAL CANCER
/ FILE REFERENCE: MRI-035
/ CURRENT APPLICATION NUMBER: US/10/171,311
/ CURRENT FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: US 60/298,159
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,155
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/335,936
/ PRIOR FILING DATE: 2001-11-14
/ NUMBER OF SEQ ID NOS: 238
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 64
/ LENGTH: 2328
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-171-311-64

Alignment Scores:
Pred. No.: 1,66e-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: Gaps: 4

US-09-940-235-11 (1-1782) x US-10-171-311-64 (1-2328)

QY 254 CCCATAGCTGAGAGTGTTCATCATGCTGCTGGAGCTTCTATGTGTCGAGAAACG 313
Db 154 Protlealaglulycyspheasphsalaalaglythrseryvalvalglulthr 173
QY 314 TGGGA-----GAAGCAGC 327
Db 174 TrrglulysprotyrclnglytrpmetmetvalasrcysthrCysleuglyglulysr 193
QY 328 GAGCGCATCACTTGACCTTCTAGAAATAGATGCAACATCAGACACAGACATCTAT 387
Db 194 GlyargilethrCysThrserArgasnArgCysasnserglinspThrargThrserTy 213
QY 388 AGAATTGAGACACCTGAGCAAGAGATTAATCGAGAAACCTGCTCGATGCAATGTC 447
Db 214 ArgileglyasprThrtrpserlylsaspsasnArglyAsnleuenglncysileCys 233
QY 448 ACAAGCAACGCGCGAGAGAGAGTGAGAGTGTGAGAGGACACCTCTGTGCAACCATG 507
Db 234 ThrGlyasnGlyArgGlyGlyutrrplysCysgluArgHsThrserValGlnThrThrser 253
QY 508 AGCGGATCTGGCCCTTCAACCGATGTCGATTGCT----- 543
Db 254 SerGlyserGlyProphethrAspValArgAlaValTyrGlnProGlnProHsPro 273
QY 544 -----GACCT 549
Db 274 GlnProProtyrGlyHisCysValThrAspserGlyValValTyrserValGlyMet 293
QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACACAGCCAAATGGTTGTTAGCTGCTGT 609
Db 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysleugly 310
QY 610 ACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATTCGATCAACATCA 669
Db 311 AsnGlyValserCysGlnGlnThrAlaValThrGlnThrTyrGlyAsnleuasnGly 330
QY 670 GACCT-----GCTCATGAGAGAAAGCA 693
Db 331 GluProCysValleuProphethrTyrAsnGlyArgThr 343

RESULT 10
US-10-236-031B-70
```

```

; Sequence 70, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Rodrick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,369
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-031B-70

```

```

Alignment Scores:
Pred. No.: 1,66e-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 4 Gaps: 4
US-09-940-235-11 (1-1782) x US-10-236-031B-70 (1-2328)

```

```

QY 254 CCCATGCTGGAAGAGTTTGTGATCATGCTGCGGACCTTCCTATGCTGGAGAAACG 313
DB 154 ProllealaglulysCyspheasphlsalalagllyThrserYValValGlyGlnThr 173
QY 314 TGGGA-----GAAGGAGC 327
DB 174 TrpGlnlyserProtyrGlnGlyTTrpMetMetValAspCysThrCybleuGlyGlnGlySer 193
QY 328 GGAAGCATCATCTTCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 387
DB 194 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 213
QY 388 AGAATTGAGACACCTGAGACAAAGAAATATCGAGAAACCTGCTCCAGTGCAATGTC 447
DB 214 ArgIleGlyAspThrTrpSerTyrLysAspAsnArgGlyAsnLeuGlnCysIleCys 233
QY 448 ACAAGCAAGGCGGAGAGAGTGAAGTGAAGGACACACCTCTGTGACACCATCG 507
DB 234 ThrGlyAsnGlyArgGlyGlnTyrLysCysGlnArgGlnThrSerValGlnThrThrSer 253
QY 508 AGCGATCTGGCCCTTCAACCGATGTTCCGATTGCT-----543
DB 254 SerGlySerGlyProphethrAspValArgAlaValTyrGlnProGlnProHisPro 273
QY 544 -----GAGACT 549
DB 274 GlnProProProtyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 293
QY 550 GAGTGGCTGCTAGACCGTCCATCTGTCAACAACCAATGCTGTTAGCGTTCGTGT 609
DB 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 310
QY 610 ACTGTTGAGGGGAGCAATCAAGACATTAGCTTAATTTTGAATTCGATCAATCAATCA 669
DB 311 AsnGlyValSerCysGlnGlnTThrAlaValThrGlnThrTyrGlyGlnAsnLeuAsnGly 330
QY 670 CGACCT-----GCTCATGGAGGAAGACA 693
DB 331 GlnProCysValLeuProphethrTyrArgGlnGlyArgThr 343

```

```

RESULT 11
US-10-374-979-98
; Sequence 98, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-98

```

```

Alignment Scores:
Pred. No.: 1,66e-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 4 Gaps: 4
US-09-940-235-11 (1-1782) x US-10-374-979-98 (1-2328)

```

```

QY 254 CCCATGCTGGAAGAGTTTGTGATCATGCTGCGGACCTTCCTATGCTGGAGAAACG 313
DB 154 ProllealaglulysCyspheasphlsalalagllyThrserYValValGlyGlnThr 173
QY 314 TGGGA-----GAAGGAGC 327
DB 174 TrpGlnlyserProtyrGlnGlyTTrpMetMetValAspCysThrCybleuGlyGlnGlySer 193
QY 328 GGAAGCATCATCTTCACTTCTAGAAATAGATGCAACGATGAGACACAAAGACATCTTAT 387
DB 194 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 213
QY 388 AGAATTGAGACACCTGAGACAAAGAAATATCGAGAAACCTGCTCCAGTGCAATGTC 447
DB 214 ArgIleGlyAspThrTrpSerTyrLysAspAsnArgGlyAsnLeuGlnCysIleCys 233
QY 448 ACAAGCAAGGCGGAGAGTGAAGTGAAGGACACACCTCTGTGACACCATCG 507
DB 234 ThrGlyAsnGlyArgGlyGlnTyrLysCysGlnArgGlnThrSerValGlnThrThrSer 253
QY 508 AGCGATCTGGCCCTTCAACCGATGTTCCGATTGCT-----543
DB 254 SerGlySerGlyProphethrAspValArgAlaValTyrGlnProGlnProHisPro 273
QY 544 -----GAGACT 549
DB 274 GlnProProProtyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 293
QY 550 GAGTGGCTGCTAGACCGTCCATCTGTCAACAACCAATGCTGTTAGCGTTCGTGT 609
DB 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 310
QY 610 ACTGTTGAGGGGAGCAATCAAGACATTAGCTTAATTTTGAATTCGATCAATCAATCA 669
DB 311 AsnGlyValSerCysGlnGlnTThrAlaValThrGlnThrTyrGlyGlnAsnLeuAsnGly 330

```

QY 670 CGACCT-----GCTCATGAGGAAAGACA 693
 Db 331 GluProCyValLeuProPheThrTyraAsnGlyArgThr 343

RESULT 12

US-10-182-936A-98
 ; Sequence 98, Application US/10182936A
 ; Publication No. US20040038860A1
 GENERAL INFORMATION:
 ; APPLICANT: Allen, Kristina M.
 ; APPLICANT: Anisowicz, Anthony
 ; APPLICANT: Bhat, Bheem
 ; APPLICANT: Damagnez, Veronique
 ; APPLICANT: Robinson, John
 ; APPLICANT: Yaworsky, Paul
 TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
 ; FILE REFERENCE: 032796-143
 ; CURRENT APPLICATION NUMBER: US/10/182,936A
 ; PRIOR FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: PCT/US02/15982
 ; PRIOR FILING DATE: 2002-05-17
 ; PRIOR APPLICATION NUMBER: US 60/291,311
 ; PRIOR FILING DATE: 2001-05-17
 ; PRIOR APPLICATION NUMBER: US 60/353,058
 ; PRIOR FILING DATE: 2002-02-01
 ; PRIOR APPLICATION NUMBER: US 60/361,293
 ; PRIOR FILING DATE: 2002-03-04
 ; NUMBER OF SEQ ID NOS: 216
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 98
 ; LENGTH: 2328
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-182-936A-98

Alignment Scores:

Pred. No.: 1,666-37 Length: 2328
 Score: 504.50 Matches: 105
 Percent Similarity: 60.10% Conservative: 11
 Best Local Similarity: 54.40% Mismatches: 28
 Query Match: 16.10% Indels: 50
 DB: Gaps: 4

US-09-940-235-11 (1-1782) x US-10-182-936A-98 (1-2328)

QY 254 CCCATAGCTGAGAAAGTGTTCATGATCATGCTGGGACTTCCTATGTGTCGAGAAACG 313
 Db 154 ProilealaglulvCyserPheAspHisAlaIaIaGlyThrSerTyraValGlyGluThr 173
 QY 314 TGGGA-----GAGGCAGC 327
 Db 174 TrpGluLysProTyrgInGlyTrpMetMetValAspCyThrCysLeuGlyGluGlySer 193
 QY 328 GAGCGATCACTTGACTTCTAGAAATAGATGCAACGATCAGAGACAAGCATCTAT 387
 Db 194 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 213
 QY 388 AGAATGGAGACACTGAGCAAGAAAGATTAATCGAGAAACCTGCTCCAGTCATCTGC 447
 Db 214 ArgIleGlyAspThrTrpSerTyrsLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 233
 QY 448 ACAGGCAACGCGCCGAGAGAGTGAAGTGTGAGGACACACTCTGTGCAACCATCG 507
 Db 234 ThrGlyAsnGlyArgGlyGlyTrpLysCysGluArgHisThrSerValGlnThrThrSer 253
 QY 508 AGCGATCGGCGCCCTTACCGATGTCGTAATGCT-----GAGCCT 543
 Db 254 SerGlySerGlyProPheThrAspValArgAlaIaIaValTyrgInProGlnProHisPro 273
 QY 544 -----GAGCCT 549
 Db 274 GlnProProProTyrgIlyHisCysValThrAspSerGlyValValTyrservalGlyMet 293

QY 550 GAGTGGCTGTAGACCGTCATCTGTCAACACAGCCATTTGGTGTACGCTGTGT 609
 Db 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCystrCysLeuGly 310
 QY 610 ACTGTTAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATCATCAATCAATCA 669
 Db 311 AsnGlyValSerCysGlnGluThrAlaValThrGlnThrTyrgIlyGlyAsnLeuAsnGly 330
 QY 670 CGACCT-----GCTCATGAGGAAAGACA 693
 Db 331 GluProCyValLeuProPheThrTyraAsnGlyArgThr 343

RESULT 13

US-10-477-238A-677
 ; Sequence 677, Application US/10477238A
 ; Publication No. US20040221326A1
 GENERAL INFORMATION:
 ; APPLICANT: Babi, Philip
 ; APPLICANT: Yaworsky, Paul
 ; APPLICANT: Bex, Frederick J. III
 ; APPLICANT: Bodine, Peter Van Nest
 TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
 ; FILE REFERENCE: 032796-212
 ; CURRENT APPLICATION NUMBER: US/10/477,238A
 ; PRIOR FILING DATE: 2003-11-10
 ; PRIOR APPLICATION NUMBER: US 60/290,071
 ; PRIOR FILING DATE: 2001-05-11
 ; PRIOR APPLICATION NUMBER: US 60/291,311
 ; PRIOR FILING DATE: 2001-05-17
 ; PRIOR APPLICATION NUMBER: US 60/353,058
 ; PRIOR FILING DATE: 2002-02-01
 ; PRIOR APPLICATION NUMBER: US 60/361,293
 ; PRIOR FILING DATE: 2002-03-04
 ; NUMBER OF SEQ ID NOS: 812
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 677
 ; LENGTH: 2328
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-477-238A-677

Alignment Scores:

Pred. No.: 1,666-37 Length: 2328
 Score: 504.50 Matches: 105
 Percent Similarity: 60.10% Conservative: 11
 Best Local Similarity: 54.40% Mismatches: 28
 Query Match: 16.10% Indels: 50
 DB: Gaps: 4

US-09-940-235-11 (1-1782) x US-10-477-238A-677 (1-2328)

QY 254 CCCATAGCTGAGAAAGTGTTCATGATCATGCTGGGACTTCCTATGTGTCGAGAAACG 313
 Db 154 ProilealaglulvCyserPheAspHisAlaIaIaGlyThrSerTyraValGlyGluThr 173
 QY 314 TGGGA-----GAGGCAGC 327
 Db 174 TrpGluLysProTyrgInGlyTrpMetMetValAspCyThrCysLeuGlyGluGlySer 193
 QY 328 GAGCGATCACTTGACTTCTAGAAATAGATGCAACGATCAGAGACAAGCATCTAT 387
 Db 194 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 213
 QY 388 AGAATGGAGACACTGAGCAAGAAAGATTAATCGAGAAACCTGCTCCAGTCATCTGC 447
 Db 214 ArgIleGlyAspThrTrpSerTyrsLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 233
 QY 448 ACAGGCAACGCGCCGAGAGAGTGAAGTGTGAGGACACACTCTGTGCAACCATCG 507
 Db 234 ThrGlyAsnGlyArgGlyGlyTrpLysCysGluArgHisThrSerValGlnThrThrSer 253
 QY 508 AGCGATCGGCGCCCTTACCGATGTCGTAATGCT-----GAGCCT 543

```
Db 254 SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 273
Qy 544 -----GACCT 549
Db 274 GlnProProTyrGlyHisCysValThrAspSerGlyValTyrSerValGlyMet 293
Qy 550 GAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGGTGGCTGTGCT 609
Db 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 310
Qy 610 ACTGTTGAGGGAGCAAGATCAAGACATTAGCTTAATTTTGAATCCGATCTAACATCA 669
Db 311 AsnGlyValSerCysGlnGlnThrAlaValThrGlnThrTyrGlyGlyAsnLeuAsnGly 330
Qy 670 CGACCT-----GCTCATGGAGAGAAAGACA 693
Db 331 GluProCysValLeuProPheThrTyrAsnGlyArgThr 343

RESULT 14
US-10-680-287A-677
; Sequence 677, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: BabiJ, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-680-287A-677

Alignment Scores:
Pred. No.: 1,666-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 5 Gaps: 4

US-09-940-235-11 (1-1782) x US-10-680-287A-677 (1-2328)
Qy 254 CCCATAGCTGAGAGTGTATTGATCATGTGCTGGAGACTTCATGNGTGGAGAAAG 313
Db 154 ProtlehlaGluwycySphneaphlaAlaGlyThrseryValValGlyGluThr 173
Qy 314 TGGGA-----GAAGCAGC 327
Db 174 TrpGluwycProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 193
Qy 328 GGAAGCATCTTTCGACTTCTAGAAATTAATGCAAGATGAGACACAGACATCTTAT 387
Db 194 GlyAygIleThrCysThrSerArGaAsnArgCyAsAsnAspGlnAspThrArGThrSerTyr 213
Qy 388 AGAATTGGAGACACCTGGAGCAAGAGATTAATCGAGAAACCTGCTCCAGTGCATCTGC 447
```

```
Db 214 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 233
Qy 448 ACAGCAACGGCCGAGAGAGTGAAGGTGAGAGCAACCTCTGTGACGACCATCG 507
Db 234 ThrGlyAsnGlyArgGlyGluTrpLysCysGluAghIsthSerValGlnThrSer 253
Qy 508 AGCGGATCTGAGCCCTTCACCGAGTTCGTATTTGCT----- 543
Db 254 SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 273
Qy 544 -----GACCT 549
Db 274 GlnProProTyrGlyHisCysValThrAspSerGlyValTyrSerValGlyMet 293
Qy 550 GAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGGTGGCTGTGCT 609
Db 294 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 310
Qy 610 ACTGTTGAGGGAGCAAGATCAAGACATTAGCTTAATTTTGAATCCGATCTAACATCA 669
Db 311 AsnGlyValSerCysGlnGlnThrAlaValThrGlnThrTyrGlyGlyAsnLeuAsnGly 330
Qy 670 CGACCT-----GCTCATGGAGAGAAAGACA 693
Db 331 GluProCysValLeuProPheThrTyrAsnGlyArgThr 343

RESULT 15
US-10-477-173-677
; Sequence 677, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
; APPLICANT: Genome Therapeutics Corporation and
; APPLICANT: Allen, Kristina M.
; APPLICANT: Yaworsky, Paul
; APPLICANT: Morales, Arturo J.
; APPLICANT: Graham, James R.
; APPLICANT: Antosiewicz, Anthony
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
; FILE REFERENCE: 032796-135
; CURRENT APPLICATION NUMBER: US/10/477,173
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-477-173-677

Alignment Scores:
Pred. No.: 1,666-37 Length: 2328
Score: 504.50 Matches: 105
Percent Similarity: 60.10% Conservative: 11
Best Local Similarity: 54.40% Mismatches: 28
Query Match: 16.10% Indels: 50
DB: 5 Gaps: 4

US-09-940-235-11 (1-1782) x US-10-477-173-677 (1-2328)
Qy 254 CCCATAGCTGAGAGTGTATTGATCATGTGCTGGAGACTTCATGNGTGGAGAAAG 313
Db 154 ProtlehlaGluwycySphneaphlaAlaGlyThrseryValValGlyGluThr 173
Qy 314 TGGGA-----GAAGCAGC 327
```

```
Db      174  TPGluLysProTyrGlnGlyTTrpMetMetValAspCysThrCysLeuGlyGluGlySer 193
      |||
QY      328  GGACGCATCACTTGCACTTCTAGAAATATGATGCAACGATCAGACACAGAACAATCCTAT 387
      |||
Db      194  GLYArgLLeThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 213
      |||
QY      388  AGAATTGAGAGACACCTGAGCAAGAAGGATTAATCGAGAAACCTGCTCCAGTCATCTGC 447
      |||
Db      214  ArgLLeGlyAspThrTTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysLLeCys 233
      |||
QY      448  ACAGGCAACGCGCAGAGAGAGTGAAGTGTGAGAGGCACACTCTGTGACAGACCACTCG 507
      |||
Db      234  ThrGlyAsnGlyArgGlyLysLysLysCysGlyLysArgHisThrSerValGlnThrThrSer 253
      |||
QY      508  AGCGGATCTGGCCCTTCAACCGATGTCGATTGCT----- 543
      |||
Db      254  SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 273
      |||
QY      544  -----GACCT 549
      |||
Db      274  GlnProProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 293
      |||
QY      550  GATGGCTGCTAGACCGTCATCTGTCAACAACAGCCAAATGSGTGTGATGCTGCTGCT 609
      |||
Db      294  GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 310
      |||
QY      610  ACTGTTGAGGGGAGCAAGATCAAGACATTAGTCTTAATTTTGAATTCGATCTACATCA 669
      |||
Db      311  AsnGlyValSerCysGlnGlnThrAlaValThrGlnThrTyrGlyGlyAsnLeuAsnGly 330
      |||
QY      670  CGACCT-----GCTCATGAGAGAAAGCA 693
      |||
Db      331  GlnProCysValLeuProPheThrTyrAsnGlyArgThr 343
      |||
```

Search completed: January 28, 2006, 02:57:14
Job time : 140.535 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 02:09:11 ; Search time 5.15975 Seconds

(without alignments)
7479.985 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134
Sequence: 1 tcgcctcacgctcgcgcg.....ataccctaagcagaataa 1782

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 75621 seqs, 10823074 residues

Total number of hits satisfying chosen parameters: 151242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DBV-xlp
-O=/cgn2_1/USFTO_spool_p/US09940235/runat_27012006_144220_27676/app_query.fasta_1.7708
-DB=Published Applications AA.New -OPMT=faetan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USRR=US09940235 @CCN_1_1 @runat_27012006_144220_27676
-LONELOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.New.*

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2158	68.9	414	6	US-10-631-558-2
2	2090.5	66.7	413	7	US-11-032-951-12
3	503	16.0	642	6	US-10-995-561-631
4	503	16.0	657	6	US-10-995-561-622
5	503	16.0	984	6	US-10-995-561-629
6	503	16.0	2296	6	US-10-995-561-633
7	503	16.0	2355	6	US-10-995-561-623
8	503	16.0	2355	6	US-10-995-561-627
9	503	16.0	2384	6	US-10-821-234-1545
10	503	16.0	2386	6	US-10-995-561-626

11	500.5	16.0	300	7	US-11-006-119-31	Sequence 31, App1
12	497.5	15.9	259	6	US-10-631-558-4	Sequence 4, App1
13	166.5	5.3	693	6	US-10-995-561-632	Sequence 632, App1
14	166.5	5.3	1259	6	US-10-995-561-625	Sequence 625, App
15	166.5	5.3	1315	6	US-10-995-561-630	Sequence 630, App
16	166.5	5.3	1341	6	US-10-995-561-621	Sequence 621, App
17	162.5	5.2	1348	6	US-10-995-561-624	Sequence 624, App
18	132.5	4.2	1992	7	US-11-013-759-13	Sequence 3, App1
19	132.5	4.2	2047	7	US-11-013-759-13	Sequence 13, App1
20	132.5	4.2	2047	7	US-11-013-759-13	Sequence 4, App1
21	132.5	4.2	2047	7	US-11-013-759-7	Sequence 7, App1
22	125	4.0	1417	7	US-11-052-554A-8	Sequence 8, App1
23	123	3.9	847	6	US-10-995-561-634	Sequence 634, App
24	122	3.9	1286	6	US-10-995-561-628	Sequence 628, App
25	116.5	3.7	470	6	US-10-485-517-319	Sequence 319, App
26	115	3.7	21	7	US-11-032-951-7	Sequence 7, App1
27	115	3.7	463	6	US-10-793-626-960	Sequence 960, App
28	115	3.7	1765	6	US-10-055-877-140	Sequence 140, App
29	114.5	3.7	1940	6	US-10-055-877-141	Sequence 141, App
30	111	3.5	801	7	US-11-174-150-29	Sequence 29, App1
31	111	3.5	801	7	US-11-124-368A-292	Sequence 292, App
32	109.5	3.5	2340	7	US-11-052-554A-171	Sequence 171, App
33	109	3.5	876	7	US-11-077-550-82	Sequence 82, App1
34	109	3.5	876	7	US-11-077-550-106	Sequence 106, App
35	109	3.5	876	7	US-11-077-550-108	Sequence 108, App
36	108.5	3.5	2314	7	US-11-013-759-11	Sequence 11, App1
37	108.5	3.5	2890	7	US-11-115-639-31	Sequence 31, App1
38	108.5	3.5	2890	7	US-11-115-639-32	Sequence 32, App1
39	108.5	3.5	2890	7	US-11-115-639-33	Sequence 33, App1
40	107	3.4	826	6	US-10-793-626-1066	Sequence 1066, App
41	106	3.4	865	6	US-10-467-962B-33	Sequence 33, App1
42	105.5	3.4	1014	7	US-11-204-187-20	Sequence 20, App1
43	105	3.4	876	7	US-11-077-550-128	Sequence 128, App
44	105	3.4	877	7	US-11-077-550-126	Sequence 126, App
45	105	3.4	877	7	US-11-077-550-130	Sequence 130, App

ALIGNMENTS

RESULT 1
US-10-631-558-2
Sequence 2, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Garish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammarra
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
CURRENT FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 414
TYPE: PRT
ORGANISM: Streptococcus equisimilis
US-10-631-558-2
Alignment Scores:

Pred. No.: 6.3e-157 Length: 414
 Score: 2158.00 Matches: 414
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.86% Indels: 0
 DB: 6 Gaps: 0

US-09-940-235-11 (1-1782) x US-10-631-558-2 (1-414)

QY	538	ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCCAATTGGTTGT	597
DB	1	llealaglProglutlrpleuLeuaspArgProSerValasnbnserGlnleuVal	20
QY	598	AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC	657
DB	21	ServalAlaaglYthrValGlnGlnGlnAspIleSerleuYsPhepegIuile	40
QY	658	GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA	717
DB	41	AspleuthrSerArgProIahlsGlyGlyThrGlnGlnGlyleuSerProlySer	60
QY	718	AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA	777
DB	61	LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGlnLysAlaAspleuLeu	80
QY	778	AAGGCTATTCAGAACAAATGATCGCTACCGTCCACAGTACAGACGACTTACTTGAGCTC	837
DB	81	LysAlaAlaGlnGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspIyrPheGlnVal	100
QY	838	ATTGATTTTGCAGCGATGACCACTTACTGATCGAAAGCGACAGGCTACTTGTGTCAC	897
DB	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValYrPheAlaAsp	120
QY	898	AAAGATGTTGCGTAACTTGGCCGACCCGACCTGTCCAGAAATTTTGTAAAGCGACAT	957
DB	121	LysAspGlySerValThrleuProThrGlnProValGlnGlnPheleuLeuSerGlyHis	140
QY	958	GTCGGCGCTTACAGCATATTAAGAAAAACCAATACAAAACCAAGCGAAATCTGTATGTCG	1017
DB	141	ValArgValArgProIyrGlyGlySerProIleGlnAsnGlnAlaLysSerValAspVal	160
QY	1018	GAATATATCTGATCAGATTACTCCCTTAAACCTGATGACGATTTCAGACAGGTCGAA	1077
DB	161	GlnYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuYs	180
QY	1078	GATACTAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGATTACTA	1137
DB	181	AspThrLysLeuLeuLysThrleuAlaIleGlyAspThrIleThrSerGlnGlnleuLeu	200
QY	1138	GCTCAGACGAAAGATTTTAAACAAAACCAACCGCTATACGATTTATGACCTGAC	1197
DB	201	AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyYrThrIleYrGlnArgAsp	220
QY	1198	TCCCTCAATCGTCACTCATGACAAATGATTTCCGATGATTTTACATGATGATCAAG	1257
DB	221	SerSerIleValIleHisAspAsnAspIlePheArgThrIleleuProMetSerGlnGln	240
QY	1258	TTTACTTACCGTGTTAAATAATCGGAAACAAGCTTATAGGATCAATAAATCTGGTCTG	1317
DB	241	PheThrYrArgValLysAsnArgGlnGlnAlaIyrArgIleAsnLysSerGlyLeu	260
QY	1318	AATGAGAAATTAACACACATGACCTGATCTCTGAGAAATATTACTCTTAAAAAGG	1377
DB	261	AsnGlnGlnLysAsnHisAspAsnThrAspLeuIleSerGlnYrYrValleuYsGly	280
QY	1378	GAAGAAGCGTATGATCCCTTGAATGCGAGTCACTGAACTGTCCATCCAAATCGTT	1437
DB	281	GlnLysProIyrAspProPheAspArgSerHisLeuYsLeuPheThrIleYrVal	300
QY	1438	GATGCTGATACCAAGAAATTTGCTAAAGTAGAGAGCTTAAACAGCTACGCAACGTAAC	1497
DB	301	AspValAspThrAsnGlnleuLeuYsSerGlnGlnleuLeuThrAlaSerGlnArgAsn	320

QY 1498 TTAGACTTGAGAGATTTATAGATCCGTGATTAAGGCTAAACTCTCAACAACATCTC 1557
 DB 321 LeuAspPheArgAspLeuYrAspProArgAspLysAlaLysleuLeuYrAsnAsnleu 340
 QY 1558 GATGCTTTGGTATTATAGACTATACCTTAACCTGAAAGTAGAGATTAATCAGATGAC 1617
 DB 341 AspAlaPheGlyIleMetAspYrThrLeuThrGlyYsValGlnAspAsnHisAspAsp 360
 QY 1618 ACCAAGCGTATCATTAACCGTTTATATGGCGACGACCCGAGAGAGAGATGCTACTAT 1677
 DB 361 ThrAsnArgIleIleThrValYrMetGlyLysArgProGlnGlnGlnValIyrSerYrleu 380
 QY 1678 CATTTAGCCTATGATTAAGTCGTTATATACCGAAGAAACAGAGAGTTTACGACTACTG 1737
 DB 381 HisleuAlaYrAspLysAspArgYrThrGlnGlnGlnGlnGlnValIyrSerYrleu 400
 QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAA 1779
 DB 401 ArgYrThrGlyThrProIleProAspAsnProAsnAspLys 414

RESULT 2
 US-11-032-951-12
 ; Sequence 12, Application US/11032951
 ; Publication No. US20050282741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kyrtal, Gerald
 ; APPLICANT: Radkin, Simon W.
 ; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
 ; FILE REFERENCE: 50216/003005
 ; CURRENT APPLICATION NUMBER: US/11/032,951
 ; PRIOR FILING DATE: 2005-01-11
 ; PRIOR APPLICATION NUMBER: US 09/919,703
 ; PRIOR FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 09/294,457
 ; PRIOR FILING DATE: 1999-04-19
 ; PRIOR APPLICATION NUMBER: US 08/759,599
 ; PRIOR FILING DATE: 1996-12-05
 ; PRIOR APPLICATION NUMBER: US 60/008,233
 ; PRIOR FILING DATE: 1995-12-06
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 413
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic polypeptide
 US-11-032-951-12

Alignment Scores:
 Pred. No.: 8.47e-152 Length: 413
 Score: 2090.50 Matches: 408
 Percent Similarity: 98.80% Conservative: 2
 Best Local Similarity: 98.31% Mismatches: 2
 Query Match: 66.70% Indels: 3
 DB: 7 Gaps: 3

US-09-940-235-11 (1-1782) x US-11-032-951-12 (1-413)

QY	538	ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCCAATTGGTTGT	597
DB	1	llealaglProglutlrpleuLeuaspArgProSerValasnbnserGlnleuVal	20
QY	598	AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC	657
DB	21	ServalAlaaglYthrVal---GlyThrAsnGlnAspIleSerleuYsPhepegIuile	39
QY	658	GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA	717
DB	40	AspleuthrSerArgProIahlsGlyGlyThrGlnGlnGlnGlnGlnleuSerProlySer	59
QY	718	AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA	777

```

Db      |||
60      LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGluValAlaAspLeuLeu 79
QY      778 AAGGCTATTCAAGAACAAATGATCGCTAACGCTCAACAGCAAGCTCTTGAAGTC 837
Db      80      LysAlaAlaLeuGlnGlnLeuLeuLeuAlaAsnValHisSerAsnAspArgTyrrPheGlnVal 99
QY      838 ATTGATTTTGCAGACGATGCAACCATTAATGATGCAAAAGCGCAAGTCTTCTTGTGTCAC 897
Db      100      IleAspPheAlaSerAspAlaThrIleThrAspArgAsnIlyValTyrrPheAlaAsp 119
QY      898 AAGATGCTTCGCTAACCTTGGCCGACCCAACTGCTCCAGAAATTTTGTCAAGCGCAAT 957
Db      120      LysAspGlySerValThrLeuProThrGlnProValGlnIlePheLeuLeuSerGlyHis 139
QY      958 GTGGCGGCTTAAGACATTAATAAGAAAACCAATACAAACCAAGCAAAATCTGTGATGTCG 1017
Db      140      ValArgValArgProTyrrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 159
QY      1018 GAATATAGTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACGCTCAAA 1077
Db      160      GluTyrrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 179
QY      1078 GATPACTAACCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGATTAATACTA 1137
Db      180      LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 199
QY      1138 GCTCAAGCAACAAAGCAATTTTAAACAAAAACACCCGAGCTATACGATTTTGAACGTGAC 1197
Db      200      AlaGlnAlaGlnSerIleLeuLeuLysAsnHisProGlyTyrrThrIleTyrrGluArgAsp 219
QY      1198 TCTCGATGCTGCTGATGACATGATCTTTCGATGCAATTTTCAAGGATGCAAG 1257
Db      220      SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleuProMetAspGlnGlu 239
QY      1258 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGATCAATTAATAATCGTGTG 1317
Db      240      PheThrTyrrArgValLysAsnArgGlnGlnAlaIleTyrrArgIleAsnLysLysSerGlyLeu 259
QY      1318 AATGAAGAAATAAACCAACACTGATCTGTGAGAAATAT--TACGCTCTTAAAAA 1374
Db      260      AsnGlnGlnIleLeuAsnThrAspLeuIleSerLeuGlnTyrrLysTyrrValLeuLys 279
QY      1375 GGGGAAAACCCGATATGATCCCTTGAATCGGACGCTGAAACTGTTCACCAATCAATAC 1434
Db      280      GlyGluLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrr 299
QY      1435 GTTGATGTGATACCAAGCAATGTGTAATAAGAGAGAGCTTTAACAGCTAGGAAAGCT 1494
Db      300      ValAspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArg 319
QY      1495 AACTTAAGATTCAGAGATTTATTCGATCCTCGTATTAAGGCTTAACTACTTCAACAAAT 1554
Db      320      AsnLeuAspPheArgLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsn 339
QY      1555 CTCGATGCTTTTGGTATTTATGCACTTACTTACGAAAGAGAAAGTAAATCAACGAT 1614
Db      340      LeuAspAlaPheGlyIleMetAspTyrrThrLeuThrGlyLysValGluAspAsnHisAsp 359
QY      1615 GACACCAACCGATTCATTAACCGTTTATATGGGACAGACCGAAGAGAGAAATGCTAGC 1674
Db      360      AspThrAsnArgGlnIleThrValTyrrMetGlyLysArgProGlnGlnGluAsnAlaSer 379
QY      1675 TATCATTTTACCTATGATTAAGATCGTTTATACCGAAGAGAAAGAGAAAGTTTACAGTAC 1734
Db      380      TyrrHis---AlaTyrrAspLysAspArgTyrrThrGlnGlnGlnArgGlnValIlyrSerTyrr 398
QY      1735 CTCGCTTATACAGGAGACACCTATACCTGATTAACCTTAACAGCAAA 1779
Db      399      LeuArgTyrrThrGlyThrProIleProAspAsnProAspAspLys 413

```

RESULT 3
US-10-995-561-631

```

; Sequence 631, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; LENGTH: 642
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-995-561-631

Alignment Scores:
Pred. No.: 3,77e-31 Length: 642
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: 6 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-631 (1-642)
QY      254 CCCATAGCTGAGAAAGTGTGATCATGCTGCGAATTCCTTAGTGTCGAGAAACG 313
Db      181 ProIleAlaGlnLysCyPheAspHisAlaIleGlyThrSerTyrrValValGlyGlnThr 200
QY      314 TGGGA-----GAAAGCAGC 327
Db      201 TrpGluLysProTyrrGlnGlyTyrrPheMetValAspCysThrCysLeuGlyGlnGlySer 220
QY      328 GAGCGCATCATCTGACCTGCTAGAAATAGATGACATGACAGACAGACATCCAT 387
Db      221 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrr 240
QY      388 AGAATTGAGACACCTGAGCAAGAGAAAGATATATGAGAAACCTGCTCAAGTATGTC 447
Db      241 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 260
QY      448 ACAGGCAACGCGGAGAGAGTGAAGCTGAGAGGCAACCTGTGTGACAGACATCG 507
Db      261 ThrGlyAsnGlyArgGlyGlnTrpLysCysGlnArgHisThrSerValGlnThrThrSer 280
QY      508 AGCGGATCTGGCCCTTACCAGATGTCGATTGCT-----543
Db      281 SerGlySerGlyProPheThrAspValArgAlaValTyrrGlnProGlnProHisPro 300
QY      544 -----GGAAGCT 549
Db      301 GlnProProProTyrrGlyHisCysValThrAspSerGlyValValTyrrSerValGlyMet 320
QY      550 GAGTGGCTGCTAGACGCTTCATCTGTCAACAACCAATGCTGTGTTAGCGTGTGCTGT 609
Db      321 GlnThrLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 337

```

RESULT 4
US-10-995-561-622
; Sequence 622, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702

SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 622
 LENGTH: 657
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-995-561-622

Alignment Scores:
 Pred. No.: 3,79e-31 Length: 657
 Score: 503.00 Matches: 101
 Percent Similarity: 66.25% Conservative: 5
 Best Local Similarity: 63.12% Mismatches: 10
 Query Match: 16.05% Indels: 45
 DB: 6 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-622 (1-657)

```

QY 254 CCATAGCTGAGAGAGTTTGTATCATGCTGGAGCTTCTATGTGTGGAGAAACG 313
    |||
Db 181 Proilealaglulwecyspheasph:salalaglYThrseryValValglglutThr 200
    |||
QY 314 TGGGA-----GAGGCAGC 327
    |||
Db 201 TrpglulysprotyrglnglyTrpmetwetalaspcysThrcysleuglygluglyser 220
    |||
QY 328 GAGCGCATCTTGACCTTCTAGAAATAGATCAAGATCAGGACAGACATCTTAT 387
    |||
Db 221 GlyArgIleThrcysThrseryasphargCysasnaspglnasphThrgThrseryr 240
    |||
QY 388 AGAATTGAGAGACACTGGAGCAAGAGTAATCGAGAAACCTGCTCCAGTCATCTGC 447
    |||
Db 241 ArgIleGlyAspThrTrpserlyslasphasnasrglyAsnleuenglnCysIleCys 260
    |||
QY 448 ACAAGCAACGCGCCAGAGAGAGTGAAAGTGTAGAGGCACACTCTGTGACACCATCG 507
    |||
Db 261 ThrGlyAsnGlyArgGlyGlyutryplysCysGluArgHisThrseryValGlnThrThrSer 280
    |||
QY 508 AGCGGATCTGGCCCTTACCGATTCGTAATGCT-----GACCT 543
    |||
Db 281 SerGlySerGlyProphethrAspValArgalalavalYrglnProGlnProHisPro 300
    |||
QY 544 -----GACCT 549
    |||
Db 301 GlnProProProTyrglyHisCysValThraspserGlyValValYrserValGlyMet 320
    |||
QY 550 GAGTGGCTGCTAGACCGCTCATCTGTCAACAACGCAATGTGTTAGCGTTGCTGCT 609
    |||
Db 321 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetleucysThrcysleugly 337
    |||

```

RESULT 5

US-10-995-561-629
 Sequence 629, Application US/10995561
 Publication No. US20050272054A1
 GENERAL INFORMATION:
 APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 TITLE OF INVENTION: DETECTION AND USES THEREOF
 FILE REFERENCE: CL001559
 CURRENT APPLICATION NUMBER: US/10/995,561
 CURRENT FILING DATE: 2004-11-24
 NUMBER OF SEQ ID NOS: 85702
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 629
 LENGTH: 984
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-995-561-629

Alignment Scores:
 Pred. No.: 4.05e-31 Length: 984
 Score: 503.00 Matches: 101
 Percent Similarity: 66.25% Conservative: 5

Best Local Similarity: 63.12% Mismatches: 10
 Query Match: 16.05% Indels: 45
 DB: 6 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-629 (1-984)

```

QY 254 CCATAGCTGAGAGAGTTTGTATCATGCTGGAGCTTCTATGTGTGGAGAAACG 313
    |||
Db 181 Proilealaglulwecyspheasph:salalaglYThrseryValValglglutThr 200
    |||
QY 314 TGGGA-----GAGGCAGC 327
    |||
Db 201 TrpglulysprotyrglnglyTrpmetwetalaspcysThrcysleuglygluglyser 220
    |||
QY 328 GAGCGCATCTTGACCTTCTAGAAATAGATCAAGATCAGGACAGACATCTTAT 387
    |||
Db 221 GlyArgIleThrcysThrseryasphargCysasnaspglnasphThrgThrseryr 240
    |||
QY 388 AGAATTGAGAGACACTGGAGCAAGAGTAATCGAGAAACCTGCTCCAGTCATCTGC 447
    |||
Db 241 ArgIleGlyAspThrTrpserlyslasphasnasrglyAsnleuenglnCysIleCys 260
    |||
QY 448 ACAAGCAACGCGCCAGAGAGAGTGAAAGTGTAGAGGCACACTCTGTGACACCATCG 507
    |||
Db 261 ThrGlyAsnGlyArgGlyGlyutryplysCysGluArgHisThrseryValGlnThrThrSer 280
    |||
QY 508 AGCGGATCTGGCCCTTACCGATTCGTAATGCT-----GACCT 543
    |||
Db 281 SerGlySerGlyProphethrAspValArgalalavalYrglnProGlnProHisPro 300
    |||
QY 544 -----GACCT 549
    |||
Db 301 GlnProProProTyrglyHisCysValThraspserGlyValValYrserValGlyMet 320
    |||
QY 550 GAGTGGCTGCTAGACCGCTCATCTGTCAACAACGCAATGTGTTAGCGTTGCTGCT 609
    |||
Db 321 GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetleucysThrcysleugly 337
    |||

```

RESULT 6

US-10-995-561-633
 Sequence 633, Application US/10995561
 Publication No. US20050272054A1
 GENERAL INFORMATION:
 APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 TITLE OF INVENTION: DETECTION AND USES THEREOF
 FILE REFERENCE: CL001559
 CURRENT APPLICATION NUMBER: US/10/995,561
 CURRENT FILING DATE: 2004-11-24
 NUMBER OF SEQ ID NOS: 85702
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 633
 LENGTH: 2296
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-995-561-633

Alignment Scores:
 Pred. No.: 4.67e-31 Length: 2296
 Score: 503.00 Matches: 101
 Percent Similarity: 66.25% Conservative: 5
 Best Local Similarity: 63.12% Mismatches: 10
 Query Match: 16.05% Indels: 45
 DB: 6 Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-633 (1-2296)

```

QY 254 CCATAGCTGAGAGAGTTTGTATCATGCTGGAGCTTCTATGTGTGGAGAAACG 313
    |||
Db 181 Proilealaglulwecyspheasph:salalaglYThrseryValValglglutThr 200
    |||
QY 314 TGGGA-----GAGGCAGC 327
    |||

```

```

Db      201  TTPGluLysProTyrGlnGlyTTrpMetValAspCysThrCysLeuGlyGluGlySer 220
OY      328  GGAGCATCTGCACTTCTAGAAATAGATGCAAGCATGAGCAAGACATCTAT 387
Db      221  GtYArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
OY      388  AGAATTGAGACACCTGAGCAAGAATATCGAAGAACTGCTCCATGCAATGTC 447
Db      241  ArgIleGlyAspThrTrpSerIlyAspAsnArgGlyAsnLeuLeuGlnCysIleCys 260
OY      448  ACAAGCAACGCCGAGAGAGAGTGAAGTGTGAGAGCAACCTCTGTGACAGACCATCG 507
Db      261  ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 280
OY      508  AGCGGATCGGCCCTTCAACCGATGTTGCTATGCT----- 543
Db      281  SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 300
OY      544  -----GGACT 549
Db      301  GlnProProProTyrGlnHisCysValThrAspSerGlyValValTyrSerValGlyMet 320
OY      550  GAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATGCTGTGAGCGTTGCTGT 609
Db      321  GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 337

```

RESULT 7

```

US-10-995-561-623
; Sequence 623, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 623
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-623

```

Alignment Scores:

```

Pred. No.: 4,69e-31 Length: 2355
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: Gaps: 3

```

US-09-940-235-11 (1-1782) x US-10-995-561-623 (1-2355)

```

OY      254  CCCATAGCTGAGAGAGTGTGATGATGCTGCGGACCTTCTATGTGTCGAGAAACG 313
Db      181  ProIleAlaGluLysCysPheAspHisAlaIleGlyThrSerTyrValGlyGluThr 200
OY      314  TGGGA-----GAAAGCAGC 327
Db      201  TrpGluLysProTyrGlnGlyTTrpMetMetValAspCysThrCysLeuGlyGluGlySer 220
OY      328  GGAGCATCTGCACTTCTAGAAATAGATGCAAGCATGAGCAAGACATCTAT 387
Db      221  GtYArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
OY      388  AGAATTGAGACACCTGAGCAAGAATATCGAAGAACTGCTCCATGCAATGTC 447
Db      241  ArgIleGlyAspThrTrpSerIlyAspAsnArgGlyAsnLeuLeuGlnCysIleCys 260

```

```

OY      448  ACAAGCAACGCCGAGAGAGTGAAGTGTGAGAGCAACCTCTGTGACAGACCATCG 507
Db      261  ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 280
OY      508  AGCGGATCGGCCCTTCAACCGATGTTGCTATGCT----- 543
Db      281  SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 300
OY      544  -----GGACT 549
Db      301  GlnProProProTyrGlnHisCysValThrAspSerGlyValValTyrSerValGlyMet 320
OY      550  GAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATGCTGTGAGCGTTGCTGT 609
Db      321  GlnTrpLeu-----LysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGly 337

```

RESULT 8

```

US-10-995-561-627
; Sequence 627, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-627

```

Alignment Scores:

```

Pred. No.: 4,69e-31 Length: 2355
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: Gaps: 3

```

US-09-940-235-11 (1-1782) x US-10-995-561-627 (1-2355)

```

OY      254  CCCATAGCTGAGAGAGTGTGATGATGCTGCGGACCTTCTATGTGTCGAGAAACG 313
Db      181  ProIleAlaGluLysCysPheAspHisAlaIleGlyThrSerTyrValGlyGluThr 200
OY      314  TGGGA-----GAAAGCAGC 327
Db      201  TrpGluLysProTyrGlnGlyTTrpMetMetValAspCysThrCysLeuGlyGluGlySer 220
OY      328  GGAGCATCTGCACTTCTAGAAATAGATGCAAGCATGAGCAAGACATCTAT 387
Db      221  GtYArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
OY      388  AGAATTGAGACACCTGAGCAAGAATATCGAAGAACTGCTCCATGCAATGTC 447
Db      241  ArgIleGlyAspThrTrpSerIlyAspAsnArgGlyAsnLeuLeuGlnCysIleCys 260
OY      448  ACAAGCAACGCCGAGAGAGTGAAGTGTGAGAGCAACCTCTGTGACAGACCATCG 507
Db      261  ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 280
OY      508  AGCGGATCGGCCCTTCAACCGATGTTGCTATGCT----- 543
Db      281  SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 300
OY      544  -----GGACT 549
Db      301  GlnProProProTyrGlnHisCysValThrAspSerGlyValValTyrSerValGlyMet 320

```

```

QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACAAAGCCCAATGTTGTTAGCGTTGCTGCT 609
Db 321 GlnTPrleu-----LysThrGlnGlnYAsnLysGlnMetLeuCySerThrCysLeuGly 337

RESULT 9
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1545

Alignment Scores:
Pred. No.: 4.7e-31 Length: 2384
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: Gaps: 3

US-09-940-235-11 (1-1782) x US-10-821-234-1545 (1-2384)
QY 254 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCGAAGCTTCTATGTGTCGAGAAAGC 313
Db 210 ProtleaGlnLysCysPheAspHisAlaAlaGlyThrSerTyValValGlyGlnThr 229
QY 314 TGGGA-----GAGGCGAGC 327
Db 230 TrpGlnLysProTyGlnGlnTyTrpMetMetValAspCysThrCysLeuGlyGlnGlySer 249
QY 328 GAGCGCATCTGCTGACCTTCTAGAAATAGATGCAACGATCAGACACAAAGACATCTAT 387
Db 250 GlyArgGlnLeuThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 269
QY 388 AGAATTGAGACACCTGAGACAAAGATTAATCGAGAAACCTGCTCCAGTCGATCTGC 447
Db 270 ArgGlnGlyAspThrTrpSerTyLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 289
QY 448 ACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGAGACACATGCG 507
Db 290 ThrGlnAsnGlnYAsnGlyGlnTrpLysCysGlnArgHisThrSerValGlnThrThrSer 309
QY 508 AGCGGATCTGGCCCTTCAACGATGTCCTATTGCT----- 543
Db 310 SerGlySerGlyProPheThrAspValArgAlaAlaValTyGlnProGlnProHisPro 329
QY 544 -----GAGCCT 549
Db 330 GlnProProTyGlnYHisCysValThrAspSerGlyValValTySerValGlyMet 349
QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACAAAGCCCAATGTTGTTAGCGTTGCTGCT 609
Db 350 GlnTPrleu-----LysThrGlnGlnYAsnLysGlnMetLeuCySerThrCysLeuGly 366

RESULT 10
US-10-995-561-626
; Sequence 626, Application US/10995561

```

```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 626
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-626

Alignment Scores:
Pred. No.: 4.7e-31 Length: 2386
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: Gaps: 3

US-09-940-235-11 (1-1782) x US-10-995-561-626 (1-2386)
QY 254 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGCGAAGCTTCTATGTGTCGAGAAAGC 313
Db 181 ProtleaGlnLysCysPheAspHisAlaAlaGlyThrSerTyValValGlyGlnThr 200
QY 314 TGGGA-----GAGGCGAGC 327
Db 201 TrpGlnLysProTyGlnGlnTyTrpMetMetValAspCysThrCysLeuGlyGlnGlySer 220
QY 328 GAGCGCATCTGCTGACCTTCTAGAAATAGATGCAACGATCAGACACAAAGACATCTAT 387
Db 221 GlyArgGlnLeuThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 240
QY 388 AGAATTGAGACACCTGAGACAAAGATTAATCGAGAAACCTGCTCCAGTCGATCTGC 447
Db 241 ArgGlnGlyAspThrTrpSerTyLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 260
QY 448 ACAGGCAACGGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACACACATGCG 507
Db 261 ThrGlnAsnGlnYAsnGlyGlnTrpLysCysGlnArgHisThrSerValGlnThrThrSer 280
QY 508 AGCGGATCTGGCCCTTCAACGATGTCCTATTGCT----- 543
Db 281 SerGlySerGlyProPheThrAspValArgAlaAlaValTyGlnProGlnProHisPro 300
QY 544 -----GAGCCT 549
Db 301 GlnProProTyGlnYHisCysValThrAspSerGlyValValTySerValGlyMet 320
QY 550 GAGTGGCTGCTAGACCGTCATCTGTCAACAAAGCCCAATGTTGTTAGCGTTGCTGCT 609
Db 321 GlnTPrleu-----LysThrGlnGlnYAsnLysGlnMetLeuCySerThrCysLeuGly 337

RESULT 11
US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ndao, Momar
; APPLICANT: Ward, Brian
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: Perichon, Regis
; APPLICANT: Ciplergen Biosystems, Inc.
; TITLE OF INVENTION: Serum Biomarkers for Chagas Disease

```

FILE REFERENCE: 016866-012130US
CURRENT APPLICATION NUMBER: US/11/006,119
CURRENT FILING DATE: 2004-12-06
PRIOR APPLICATION NUMBER: US 60/527,153
PRIOR FILING DATE: 2003-12-05
PRIOR APPLICATION NUMBER: US 60/565,093
PRIOR FILING DATE: 2004-04-22
PRIOR APPLICATION NUMBER: US 60/625,519
PRIOR FILING DATE: 2004-11-06
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 300
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
US-11-006-119-31

Alignment Scores:
Pred. No.: 5,14e-31 Length: 300
Score: 500.50 Matches: 95
Percent Similarity: 84.82% Conservative: 0
Best Local Similarity: 84.82% Mismatches: 2
Query Match: 15.97% Indels: 16
DB: 7 Gaps: 1

US-09-940-235-11 (1-1782) x US-11-006-119-31 (1-300)

QY 254 CCATAGCTGAGAAAGTTTGAATCATGCTGCGGACCTTCTATGTGTGGAGAAACG 313
|||
DB 150 Protlealaglulwscyspheasphialalaglthrseryvalvalglulthr 169
QY 314 TGGGA-----GAGGCACG 327
|||
DB 170 TrpdlulwspotyrginglYTrpMetMetValaerpCysThrCysleuglYglulYser 189
QY 328 GAGCGATCATCTGCACTTCTAGAAATAGATGCAACGATGAGACACAGCATCTCTAT 387
|||
DB 190 GlYargilethrCysThrSerArghsnArgCysAsnAspGlnAspThrArgThrserYr 209
QY 388 AGAATTGAGACACCTGAGCAAGAGATATGAGAAACCTGCTCAAGTCATCTGC 447
|||
DB 210 ArgileglYasphrTrpserYslyAspAsnArglyAsnleuGlnCyslleCys 229
QY 448 AAGCGAAGCGCCGAGAGAGTGAAGTGTGAGAGGCACCTCTGTGAGACACATCG 507
|||
DB 230 ThrGlYasnGlYarglYglulTrpYsCysgluArgHsrThrserValGlnThrThrser 249

QY 508 AGCGATCTGGCCCTTCAACCGATGTCGATTGCT 543
|||
DB 250 SerGlYserGlYProPhetThrAspValArgAlaAla 261

RESULT 12
US-10-631-558-4
Sequence 4, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammaru
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
CURRENT FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235

PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 259
TYPE: PRT
ORGANISM: Homo sapiens
US-10-631-558-4

Alignment Scores:
Pred. No.: 8,48e-31 Length: 259
Score: 497.50 Matches: 94
Percent Similarity: 85.45% Conservative: 0
Best Local Similarity: 85.45% Mismatches: 1
Query Match: 15.87% Indels: 16
DB: 6 Gaps: 1

US-09-940-235-11 (1-1782) x US-10-631-558-4 (1-259)

QY 254 CCATAGCTGAGAAAGTTTGAATCATGCTGCGGACCTTCTATGTGTGGAGAAACG 313
|||
DB 150 Protlealaglulwscyspheasphialalaglthrseryvalvalglulthr 169
QY 314 TGGGA-----GAGGCACG 327
|||
DB 170 TrpdlulwspotyrginglYTrpMetMetValaerpCysThrCysleuglYglulYser 189
QY 328 GAGCGATCATCTGCACTTCTAGAAATAGATGCAACGATGAGACACAGCATCTCTAT 387
|||
DB 190 GlYargilethrCysThrSerArghsnArgCysAsnAspGlnAspThrArgThrserYr 209
QY 388 AGAATTGAGACACCTGAGCAAGAGATATGAGAAACCTGCTCAAGTCATCTGC 447
|||
DB 210 ArgileglYasphrTrpserYslyAspAsnArglyAsnleuGlnCyslleCys 229
QY 448 AAGCGAAGCGCCGAGAGAGTGAAGTGTGAGAGGCACCTCTGTGAGACACATCG 507
|||
DB 230 ThrGlYasnGlYarglYglulTrpYsCysgluArgHsrThrserValGlnThrThrser 249

QY 508 AGCGATCTGGCCCTTCAACCGATGTCGATTGCT 537
|||
DB 250 SerGlYserGlYProPhetThrAspValArg 259

RESULT 13
US-10-995-561-632
Sequence 632, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: C1001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 632
LENGTH: 693
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-632

Alignment Scores:
Pred. No.: 1,41e-05 Length: 693
Score: 166.50 Matches: 36
Percent Similarity: 44.83% Conservative: 16
Best Local Similarity: 31.03% Mismatches: 41
Query Match: 5.31% Indels: 24

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:35 ; Search time 27.3089 Seconds

(without alignments)
12556.953 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134
Sequence: 1 tcgtctcacgttcgtctgcg.....ataccctaagacaataa 1782

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
MODEL=frame+ n2p.model -DBV=xld  
-O=/cgn2.1/USPNO_spool_P/US09940235/runat_27012006_144218_27578/app_query.fasta_1.7708  
-DB=PIR -QFMT=fastan -SUFRTX=tpir -MINMATCH=0.1 -LOOPLC=0 -LOOPEXT=0  
-ONITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIG=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09940235 @CCN 1.1.185 @runat_27012006_144218_27578 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEDEUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -YGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
PIR 80:*  
1: pir1:*  
2: pir2:*  
3: pir3:*  
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2158	68.9	440	1 A22801	streptokinase prec
2	2120.5	67.7	415	1 BZSO	streptokinase A (B)
3	2116	67.5	440	2 S02723	streptokinase G pr
4	1938	61.8	440	2 S02724	streptokinase A pr
5	1841	58.7	414	2 JU0292	streptokinase A pr
6	1838	56.6	440	2 S04168	streptokinase A pr
7	788.5	25.2	197	2 S53334	streptokinase - St
8	620	19.8	128	2 S77671	streptokinase A (E)
9	569	18.2	128	2 S77680	streptokinase A (E)
10	557	17.8	128	2 S77688	streptokinase A (E)
11	552	17.6	128	2 S77679	streptokinase A (E)
12	503	16.0	2386	1 FNHU	fibronectin precu
13	481	15.3	2265	1 FNBO	fibronectin bov1
14	454.5	14.5	2477	2 S14428	fibronectin precu

15	423	13.5	128	2 S77676	streptokinase A (B)
16	420.5	13.4	2481	2 A43908	fibronectin - Afri
17	419	13.4	128	2 S77673	streptokinase A (B)
18	419	13.4	128	2 S77687	streptokinase A (B)
19	417	13.3	128	2 S77682	streptokinase A (B)
20	416	13.3	128	2 S77678	streptokinase A (B)
21	408	13.0	128	2 S77675	streptokinase A (B)
22	407	13.0	128	2 S77685	streptokinase A (B)
23	401	12.8	128	2 S77683	streptokinase A (B)
24	398	12.7	128	2 S77686	streptokinase A (B)
25	397	12.7	128	2 S77684	streptokinase A (B)
26	396	12.6	128	2 S77681	streptokinase A (B)
27	390	12.4	128	2 S77674	streptokinase A (B)
28	386	12.3	128	2 S77675	streptokinase A (B)
29	147	4.7	1020	2 A29355	fibronectin - chic
30	140.5	4.5	2231	2 D71870	hypothetical prote
31	136.5	4.4	13055	2 T16580	hypothetical prote
32	127.5	4.1	928	2 C81265	probable lipoprote
33	127.5	4.1	978	2 D81411	probable lipoprote
34	127	4.1	1965	2 S75200	fat protein - Syne
35	127	4.1	2256	2 A21018	large repetitive p
36	126	4.0	1019	2 T50251	hypothetical coile
37	125	4.0	1417	2 H90670	probable invasiv
38	125	4.0	1417	2 D85521	probable adhesin
39	125	4.0	2140	2 P95074	serine proteinase
40	124.5	4.0	190	2 I51279	hypothetical prote
41	124.5	4.0	617	2 D66978	hypothetical prote
42	124	4.0	1002	2 T02458	hypothetical prote
43	124	4.0	3394	2 T18501	hypothetical prote
44	123.5	3.9	1176	2 T47444	hypothetical prote
45	123	3.9	1286	2 T16507	hypothetical prote

ALIGNMENTS

```
RESULT 1  
A22801  
streptokinase precursor - Streptococcus "equisimilis"  
C/Species: Streptococcus "equisimilis"  
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004  
C/Accession: A22801  
R/Malke, H.; Roe, B.; Perretti, J.J.  
Gene 34, 357-362, 1985  
A/Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H  
A/Reference number: A22801, MUID:85232082; PMID:2969113  
A/Accession: A22801  
A/Molecule type: DNA  
A/Residues: 1-440 <MAL>  
A/Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:9407876; PIDN  
A/Experimental source: strain H46A  
A/Genetics:  
A/Gene: skk  
C/Superfamily: streptokinase
```

Alignment Scores:

Pred. No.:	6.23e-140	Length:	440
Score:	2158.00	Matches:	414
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.86%	Indels:	0
DB:	1	Gaps:	0

US-09-940-235-11 (1-1782) x A22801 (1-440)

QY	538	ATTGCTGACGCTGAGTGGCTGATGACGCTCATCTGTCAACAACAGCAATTGGTTGTT	597
DB	27	IIeHlaelyProGInttrpLeuLeuAspArgProSerValbnsnSerGlnLeuValVal	46
QY	598	AGCGTTGCTGTACTGTGAGGAGGAGCAATCAAGACATTAGTTAAATTTTGAATC	657
DB	47	SerValaIaelyhrValIgluGlyThrAsnGlnAspIleSerLeuIysPhePheGluIle	66
QY	658	GATCTAACATCAGCACTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA	717

```

Db      67  AspleuThrSerArgProIahIseGlyLysThrGluGlnGlyLeuSerProLysSer 86
Qy      718  AAACCATTTGCTACTGATAGTGCGCGGATGTCACATAACTTGAGAAAGCTGACTTACTA 777
Db      87  LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGlnGlyLysAlaAspLeuLeu 106
Qy      778  AAGGCTATTCAGAACAAATGATTCGCTAACGTCACAGTAAGAGACACTTGAGAGTC 837
Db      107  LysAlaAlaGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAspAspTyrPheGlnVal 126
Qy      838  ATTGATTTTGCAGAGATGCAACCACTTACTGATTCGAACGGAAGGCTTACTTGCTGAC 897
Db      127  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlyrPheAlaAsp 146
Qy      898  AAAAGATGGTTCGGTAACTTGCCGACCAACTGTCGCAAGAAATTTTGGTAAAGCGGACAT 957
Db      147  LysAspGlySerValThrLeuProThrGlnProValGlnGlnIupheLeuLeuSerGlnHis 166
Qy      958  GTGCGCGTTAGCCATATTAAGAAAAACCAATTCAAAAACCAAGCCGAAATCTGTGATGTG 1017
Db      167  ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 186
Qy      1018  GAATATCTGTAAGATTACTCCCTTAAACCTTGATGACGATTTTGACAGGCTGCAAA 1077
Db      187  GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlnLysLys 206
Qy      1078  GATACTAAGCTATTTGAAAACACTAGCTATCGGTGACACCATGACATCTCAAGAAATTTACTA 1137
Db      207  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
Qy      1138  GCTCAGACGACAAAGATTTTAAACAAAACCAACCGGCTATTCGATTTATGACGTGAC 1197
Db      227  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlnLysTyrThrIleTyrGlnLysAsp 246
Qy      1198  TCCCTCAATCGTCACCTCATGACATGACATTTTCCGATGATTTTACCAATGGATCAAGAG 1257
Db      247  SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProAspGlnGln 266
Qy      1258  TTTACTTACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG 1317
Db      267  PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlnLysLeu 286
Qy      1318  AATGAGAAATTAACAACAACCTGATCTGTGAGAAATTTATAGCTCTTAAAAAAGG 1377
Db      287  AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrLysValLeuLysLysGly 306
Qy      1378  GAAAGCGGTAATGATCCCTTGAATGCGATCACTGGAATCTGTCAACCAATACGTT 1437
Db      307  GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Qy      1438  GATGTCGATACCAAGAAATTCGTAATAAAGTGACAGCTCTTAAACAGTCCGAACTTAAC 1497
Db      327  AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysGln 346
Qy      1498  TTAAGCTTCAGAGATTTATACGATCTCGTGAATAGGCTTAAACTACTCTTACCAACATTC 1557
Db      347  LeuAspPheArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysAsnLeu 366
Qy      1558  GATGCTTTTGGTATTAATGACATTAACCTTAACTGGAAGAAAGTAAAGATTAATCAACATGAC 1617
Db      367  AspAlaPheGlyLysIleMetAspTyrThrLeuThrGlnLysValGlnAspAsnHisAspAsp 386
Qy      1618  ACCAACCTATCATTAACCGTTATATAGGCAAGCGACCGAAGAGAGAGATGCTATCTAT 1677
Db      387  ThrAsnArgIleIleThrValIlyrMetGlyLysArgProGlnGlnLysAsnAlaSerTyr 406
Qy      1678  CATTAGCTTATGATTAAGATCGTTATATACGAAGAAAGCAAGATTTTACAGCTTACCTG 1737
Db      407  HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnLysArgLysValIlyrSerTyrLeu 426
Qy      1738  CGTTATACAGGACACCTTACTGCTAATACCTTAAGACAA 1779

```

```

Db      427  ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440
RESULT 2
BZSO
Streptokinase (EC 3.4.-.-) - Streptococcus sp.
C/Species: Streptococcus sp.
C/Date: 05-Apr-1983 #sequence_rev: 05-Apr-1983 #text_change 09-Jul-2004
C/Accession: A00967
R/Jackson, K.W.; Tang, J.
Biochemistry 21, 6620-6625, 1982
A/Title: Complete amino acid sequence of streptokinase and its homology with serine protease
A/Reference number: A00967; PMID:83127125; PMID:6760891
A/Accession: A00967
A/Molecule type: protein
A/Residues: 1-415 <0AC>
A/Cross-references: UNIPROT:P00779; UNIPARC:UPI0000172BF5
A/Note: 169-Asp and 181-Asp were also found
A/Note: this protein is not a protease, but it activates plasminogen by complexing with C/superfamily: streptokinase
C/Keywords: hydrolase

Alignment Scores:
Pred. No.: 2,276-137 Length: 415
Score: 2120.50 Matches: 410
Percent Similarity: 99.28% Conservative: 2
Best Local Similarity: 98.80% Mismatches: 2
Query Match: 67.66% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-11 (1-1782) x BZSO (1-415)

Qy      538  ATTGCTGACCTGAGTGGCTGCTAAGCCGTCACATCTGTCAACAACAGCAATGGTTGTT 597
Db      1  IleAlaGlyProGlnThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
Qy      598  AGCGTTCGCTGCTGCTGTAAGGGGAGCAATCAAGACATTAAGCTTAAATTTTGAATATC 657
Db      21  SerValAlaGlyThrValGlnGlnLysThrAsnGlnAspIleSerLeuLysPhePheGlnIle 40
Qy      658  GATCTAATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAACTGCAAAAATCA 717
Db      41  AspleuThrSerArgProIahIseGlyLysThrGlnGlnGlnLysLeuSerProLysSer 60
Qy      718  AAACCATTTGCTACTGATAGTGCGCGGATGTCACATAACTTGAGAAAGCTGACTTACTA 777
Db      61  LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
Qy      778  AAAGCTATTCAGAACAAATGATTCGCTAACGTCACAGTCAACGACTTCTTGAGGTC 837
Db      81  LysAlaAlaGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100
Qy      838  ATTGATTTTGCAGAGATGCAACCACTTACTGATTCGAACGGAAGGCTTACTTGCTGAC 897
Db      101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlyrPheAlaAsp 120
Qy      898  AAAAGATGGTTCGGTAACTTGCCGACCAACTGTCGCAAGAAATTTTGGTAAAGCGGACAT 957
Db      121  LysAspGlySerValThrLeuProThrGlnProValGlnGlnIupheLeuLeuSerGlnHis 140
Qy      958  GTGCGCGTTAGCCATATTAAGAAAAACCAATTCAAAAACCAAGCGAAATCTGTGATGTG 1017
Db      141  ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
Qy      1018  GAATATCTGTAAGATTACTCCCTTAAACCTTGATGACGATTTTGACAGGCTGCAAA 1077
Db      161  GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlnLysLys 180
Qy      1078  GATACTAAGCTATTTGAAAACACTAGCTATCGGTGACACCATGACATCTCAAGAAATTTACTA 1137
Db      181  LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
Qy      1138  GCTCAGACGACAAAGATTTTAAACAAAACCAACCGGCTATTCAGATTTTATGAAAGCTGAC 1197

```

Db 201 AAGlAlaGlnSerIleuAenlysaNhiPProGlyTrhIleTyGluAAsp 220
Qy 1198 TCCTCAATGTCACATGATGACATATTCCTGATGATTTTACCATGATGATCAAG 1257
Db 221 SerSerIleValThrIrisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 1258 TTTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCGTGTG 1317
Db 241 PheThrTyArgValIlysaAsnArgGlnGlnAlaTyArgIleAenlysaSerGlyLeu 260
Qy 1318 AATGAAGAATAAACAACATGACCTGATCTTGAGAAATAT--TACGCTCTTAAAAA 1374
Db 261 AaNGluGlnIleAsnAsnThrAspLeuIleSerLeuGlnTyArgTyValIleuTyAs 280
Qy 1375 GGGGAAAGCCGATGATCCCTTGTATGCGAGTCACCTTGAAACGTTACCATCAATAC 1434
Db 281 GlyGluAspProTyArgAspProPheAspArgSerHisLeuAspLeuPheTrIleTy 300
Qy 1435 GTTGATGTGATACCAACAATTTGCTAAAAAGTGAAGAGCTTTAAACGCTAGGAAAGCT 1494
Db 301 ValAspValAspThrAsnGlnLeuLeuTySerGlnGlnLeuThrAlaSerGlnArg 320
Qy 1495 AACTTACATTCAGACATTTATTCGATCCTCGTATAGAGCTTAACTCTTACACAT 1554
Db 321 AsnLeuAspPheArgAspLeuTyArgAspProArgAspLysAlaLysLeuLeuTyAsnAsn 340
Qy 1555 CTGAGATGCTTTGGTATTATGACCTATACCTTAACTGGAAGATAGAGATATTCAGAT 1614
Db 341 LeuAspAlaPheGlyIleMetAspTyTrhLeuThrGlyLysValGluAspAsnHisAsp 360
Qy 1615 GACACCAACCGTATCAATAACCGTTTATATGAGCAAGCAAGCAAGAGAGAAATGCTAGC 1674
Db 361 AspThrAsnAspArgIleIleThrValTyMetGlyLysArgProGlnGlnAlaSer 380
Qy 1675 TATCATTTAGCTATGATTAAGATGCTTATATCCGAAAGAACGAAAGTTTACAGTAC 1734
Db 381 TyrHisLeuAlaTyArgAspAspArgTyTrhGlnGlnGlnValTySerTy 400
Qy 1735 CTGGCTTATACAGGACACCTATACCTGATTAACCTTAACGACAAA 1779
Db 401 LeuArgTyTrhGlyTrhProIleProAspAsnProAspAspLys 415
RESULT 3
S02723
Streptokinase G precursor - Streptococcus sp. (strain 19908)
C:Species: Streptococcus sp.
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02723
R:Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1262, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.
A:Reference number: S02723; MID:89160265; PMID:2922269
A:Accession: S02723
A:Molecule type: DNA
A:Residues: 1-440 <MAL>
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:947095; PIDD
C:Genetics:
A:Gene: skg
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <Sig>
F:27-440/Product: streptokinase #status predicted <Mat>
Alignment Scores:
Pred. No.: 4,64e-137 Length: 440
Score: 2116.00 Matches: 406
Percent Similarity: 98.55% Conservative: 2
Best Local Similarity: 98.07% Mismatches: 6
Query Match: 67.52% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-11 (1-1782) x S02723 (1-440)
Qy 538 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTGTTGTT 597

Db 27 IleaGlnProGlnTyPLeuLeuAspArgProSerValaAsnSerGlnLeuVala 46
Qy 598 AGCGTCCTGCTACTGTTGAGGGAGGACAAATGAAACATTAATTTTGAATC 657
Db 47 SerValaGlnTyTrhValGlnGlyTrhAsnGlnAspIleSerLeuTyPhePheGlnIle 66
Qy 658 GATCAATCAATCAACAGCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 717
Db 67 AspLeuThrSerArgProAlaHisGlyGlyTySerThrGlnGlnGlnTyLeuSerProLysSer 86
Qy 718 AAACATTTGCTACTGATAGTGGCGGATGTCACATAAATTTGAGAAAGCTGACTTACTA 777
Db 87 LysLeuPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnTyValAspLeuLeu 106
Qy 778 AAGCTTATCAAGAACATTAATGCTTAAGCTCAACGATTAACAGCATCTTGAAGTC 837
Db 107 LysAlaIleGlnGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyPheGlnVal 126
Qy 838 ATTGATTTGGACAGATGCAACATTAATGATGAAACGGCAGAGCTACTTGTGTCAC 897
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyPheAlaAsp 146
Qy 898 AAAGATGCTTCGATTAACCTTGCCGACCAACCTGTCCAAGAAATTTGCTAAGCGACAT 957
Db 147 LysAspGlySerValThrLeuProIleGlnProValGlnGlnPheLeuLeuTyGlyHis 166
Qy 958 GTGGCGCTTGAACCATTAATAAGAAAAACCAATCAAAACCAAGGAAATCTGTGATGTG 1017
Db 167 ValArgValArgProTyArgGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
Qy 1018 GAATATATCTGATACGTTTACCTCCTTAAACCTGATGACATTTCAAGCAAGCTCAAA 1077
Db 187 GluTyTrhValGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
Qy 1078 GATTAAGTATTTGAAACATTAAGTATGCTGATGACACCATCAATCATCAAGATTAATA 1137
Db 207 AspThrTyLeuLeuLeuTySerThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 226
Qy 1138 GCTCAACGACAAAGCATTTTAAACAAAAACCAACGAGCTTATACGATTTATGAACGTGAC 1197
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyTrhIleTyGluAAsp 246
Qy 1198 TCCTCAATGCTCATCATGACAAATGACATTTTCCTGATGATTTTACCAATGATGACAGAG 1257
Db 247 SerSerIleValThrIrisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy 1258 TTTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCGTGTG 1317
Db 267 PheThrTyTrhIleValIlysaAsnArgGlnGlnAlaTyArgIleAenlysaSerGlyLeu 286
Qy 1318 AATGAAGAATAAACAACATGACCTGATCTTGAGAAATATTAACGCTTAAAAAAGG 1377
Db 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyTrhValIleuTyLysGly 306
Qy 1378 GAAAAGCCGATGATCCCTTGTATGCGAGTCACCTTGAACSTGTTACCATCAATACGTT 1437
Db 307 GluTySerProTyArgAspProPheAspArgSerHisLeuTySerPheTrIleTyVal 326
Qy 1438 GATGCTATCAACAAGAAATTTGCTAAAAAGTGAAGAGCTTTAAACGCTAGGAAAGCTTAC 1497
Db 327 AspValaAsnThrAsnGlnLeuLeuTySerGlnGlnLeuThrAlaSerGlnValAsn 346
Qy 1498 TTAGACTTCAAGATTTATACGATCCCTGATTAAGGCTTAACTCTCAACATATCTC 1557
Db 347 LeuAspPheArgAspLeuTyArgAspProArgAspLysAlaLysLeuLeuTyArgAsnAsnLeu 366
Qy 1558 GATGCTTTGGTATTATGACATATACCTTAACTGAAAGTGAAGATTAATCAAGATGAC 1617
Db 367 AspAlaPheGlyIleMetAspTyTrhLeuThrGlyLysValGluAspAsnHisAspAsp 386
Qy 1618 ACCAACCGTATCAATAACCGTTTATATGAGGACGAGCAAGAGAGGATGCTGACTAT 1677

```

Db      387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluGlnAsnAlaSerTyr 406
QY      1678 CATTTAGCCTATGATAAAGATGTTATACCGAAGAAGACGAGCTTACACTGCTG 1737
Db      407 HisLeuAlaIleTyrAspLysAspArgTyrThrGluGluGlnArgGluValTyrSerTyrLeu 426
QY      1738 CGTTATACAGGAGCACCTATACCTGATTAACCTTAACGACGAAA 1779
Db      427 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440

RESULT 4
502724
Streptokinase A precursor - Streptococcus pyogenes (strain SF130/13)
C:Species: Streptococcus pyogenes
C>Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #ext_change 09-Jul-2004
C:Accession: S02724
R:Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1261, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes lys
A:Reference number: S02724; MUID:89160264; PMID:2646590
A:Accession: S02724
A:Molecule type: DNA
A:Residues: 1-440 <MAL>
C:Cross-references: UNIPROT:P10520; UNIPARC:UPI000017021D; EMBL:X13399; NID:G47435; PIDN
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <Sig>
F:27-440/Product: streptokinase #status predicted <MAT>

Alignment Scores:
Pred. No.: 6 94e-125 Length: 440
Score: 1938.00 Matches: 366
Percent Similarity: 93.72% Conservative: 22
Best Local Similarity: 88.41% Mismatches: 26
Query Match: 61.84% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x S02724 (1-440)
QY      538 ATTGCTGACCTGAGTGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGTTGTT 597
Db      27 IleAlaGlyTyrGlyTyrIleuProAspArgProIleAsnAsnSerGlnLeuVal 46
QY      598 AGCGTGTCTGACTGTTGAGGAGGAGCAATCAACAGCACTTAAATTTTGAATC 657
Db      47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnHepheGluIle 66
QY      658 GATCTAACATCAACGACCTGCTCATGAGGAAAGACGACGAGCTTAAATCCAAATCA 717
Db      67 AspIleuThrSerGlnProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 86
QY      718 AAACCATTTTGTACTGATGATGAGCGGATGTCACATTAATCTTGAGAAAGCTGACTTACTA 777
Db      87 LysProPheAlaIleThrAspAsnGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
QY      778 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACGACGACTTACTTGAAGTC 837
Db      107 LysAlaIleGlnIleGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheAlaAsp 126
QY      838 ATTGATTTTGGACGAGGATGCAACCATTTACTGATCGAAAGCGCAAGCTTACTTGTGAC 897
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspAsnGlyLysValTyrPheAlaAsp 146
QY      898 AAAGATGTTGCTGATACTTGGCGACCGAACGTCGCAAGATTTTGGCAAGCGGACAT 957
Db      147 LysAspGlySerValThrIleuProThrGlnProValGlnGlnPheLeuValLysGlyHis 166
QY      958 GTGCGCGTTTGAACCATTAAGAAAAACCAATACAAACCAAGCGCAATCTGTTGATGTC 1017
Db      167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY      1018 GAATATACTGTACAGTTTACTCCCTTAAACCCGATGACGATTTTCAGACCAAGCTCTCAAA 1077

```

```

Db      187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY      1078 GATACTAGACTATTGAAAACACTAGCTATGCTGACCAACATCACATCTCAAGAAATTACTA 1137
Db      207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleLeu 226
QY      1138 GCTCAAGCAAGAGCATTTTAAACAAACACCCAGCGTATCGATTTTGAACGCGAC 1197
Db      227 AlaGlnAlaGlnIleThrIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
QY      1198 TCCTCAATCGTCACTCATGACCAATGACATTTTCGTCGATTTTACCAAGGATCAAGAG 1257
Db      247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln 266
QY      1258 TTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGCTCTG 1317
Db      267 PheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnProLysThrGlyIle 286
QY      1318 AATGAAGAAATTAACCAACTGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGG 1377
Db      287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrValLeuLysGlnGly 306
QY      1378 GAAAGCCGATATGATCCCTTGTACGACGACCTTGAAACCTGTTACCATCAATACGTT 1437
Db      307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY      1438 GATGTCGATNACCAAGCAATGTCTTAAAGTAGACAGCTTAAACGCTTACGACGACATAC 1497
Db      327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGluGlyAsn 346
QY      1498 TTAGACTTCAGAGATTTATTCGATCCTCGTATAGGCTTAACTACTCTTCAACAATCTC 1557
Db      347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY      1558 GATGCTTTTGGTATTATGACCTATACCTTAACTGAAAGATAGAGATTAATTCAGATAC 1617
Db      367 AspAlaPheAspIleLeuAsnTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY      1618 ACCAACCGTATCATTAACCGTTTATATGGGCAACGACCGCAAGAGAGAAATGCTAGCTAT 1677
Db      387 AsnAsnArgValAlaThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY      1678 CATTTAGCCTATGATAAAGATGTTATACCGAAGAAGACGAGAGTTTACAGCTTACTG 1737
Db      407 HisLeuAlaIleTyrAspLysAspLeuTyrThrGluGluGlnArgLysAlaTyrSerTyrLeu 426
QY      1738 CGTTATACAGGAGCACCTATACCTGATTAACCTTAACGACGAAA 1779
Db      427 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440

RESULT 5
JU0292
Streptokinase - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #ext_change 09-Jul-2004
C:Accession: A43867; JU0292
R:Ohkuni, H.; Todome, Y.; Suzuki, H.; Mizuse, M.; Kotani, N.; Horiiuchi, K.; Shikama, N.;
Infect. Immun. 60, 278-283, 1992
A:Title: Immunochemical studies and complete amino acid sequence of the streptokinase fr
A:Reference number: A43867; MUID:92104686; PMID:1370275
A:Accession: A43867
A:Molecule type: protein
A:Residues: 1-414 <OHK>
C:Cross-references: UNIPROT:Q57391; UNIPARC:UPI0000175C82
A:Experimental source: M type 12 strain A374
A:Note: sequence extracted from NCBI backbone (NCBIP:74592)
C:Superfamily: streptokinase

Alignment Scores:
Pred. No.: 2 97e-118 Length: 414
Score: 1841.00 Matches: 354
Percent Similarity: 91.06% Conservative: 23

```

Best Local Similarity: 85.51% Mismatches: 37
 Query Match: 58.74% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x JU0292 (1-414)

```

QY 538 ATTGCTGACCTGAGGCTGCTGACACCGTTCATCTGTCAACAACAGCAATTGGTTGTT 597
DB 1 TlelaaglyProdiuttrpleuAspArgProSerValAsnSerGlnLeuVal 20
QY 598 AGGTTGCTGCTGCTGAGGGAACAACAAGCACTTGTCTTAATTTTGAATC 657
DB 21 SerValaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
QY 658 GATCTTAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCAAAATCA 717
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlnGlnGlnGlnGlnGln 60
QY 718 AAACCATTTGCTGATGATAGTGCGCGAGATGTCATTAACCTTGAGAAAGCTGACTTA 777
DB 61 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGlnGlnGlnGlnGlnGln 80
QY 778 AAGGCTATTCAGAAACAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
DB 81 LysAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
QY 838 ATTGATTTTGAAGCGATGCAACCATTAATGATGCAAGGCAAGGCTGCTGCTGCTGCTG 897
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyPheAlaAsp 120
QY 898 AAAGATGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
DB 121 ArgAspAspSerValThrLeuProThrGlnProValGlnGlnGlnGlnGlnGlnGlnGln 140
QY 958 GTGCGGCTTGAACCATATTAAGAAAAACCAATACAAACAGCGAAATGTTGATGTCG 1017
DB 141 ValArgValArgProGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
QY 1018 GAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
DB 161 AsnTyrgLysValSerPheValSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
QY 1078 GATCTTAAGCTTATGAAAACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
DB 181 GluArgTygHisLeuThrThrLeuAlaValGlyAspSerLeuSerGlnGlnGlnGlnGln 200
QY 1138 GCTTCAAGCAACAAGATTTTAAACAAAACACCCAGGCTATACGATTTTGAACGCTGAC 1197
DB 201 AlaIleAlaGlnPheIleLeuSerLysGlnHisProAspTyrlleIleThrLysValAsp 220
QY 1198 TCCCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1258 TTTACTTACCGTGTAAAAATGGGAAACAGCTTATAGGATCAATTAATAAATCGTGGC 1317
DB 241 PheThrTygHisIleLysAspArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
QY 1318 AATGAAGAATTAACAACATGACCTGATCTGTGAGAAATATTACGCTTAAAAAAAGG 1377
DB 261 ValGlnLysThrAsnAsnThrAspLeuIleSerGlnLysGlyTygValLeuLysGly 280
QY 1378 GAAAACCGTATGATCCCTTTGATGCGAGTCACTTGAATCTGTCACCATCAATACGTT 1437
DB 281 GlnGlnProTygAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTygVal 300
QY 1438 GATTCGATACCAAGAAATTTGCTAAATAATGACGACCTTTAACAGCTAGCAAGCTAAC 1497
DB 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
QY 1498 TTAGACTTCAGAGATTTTATACGATCTCGGATTAAGGCTTAATACCTTACCAACATTC 1557
DB 321 LeuAspPheAspGlnAspLeuTygAspProAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 340
  
```

```

QY 1558 GATGCTTTGGTATTATGACCTTAACTGGAAGTAGAGATTAATCAGATGAC 1617
DB 341 AspAlaPheGlyIleMetAspTyrlThrLeuThrGlyLysValGlnAspAsnHisAsnAsp 360
QY 1618 ACCAACCGTATCATTAACCGTTTATATGAGGCAAGCAGCCGAGAGAGAAATGCTAT 1677
DB 361 ThrAsnArgIleIleThrValTygMetGlyLysArgProGlnGlnGlnGlnGlnGlnGln 380
QY 1678 CATTAGCTTATGATTAAGATCGTTTATCCGAAAGAAAGAGAAAGTTTACGCTACCTG 1737
DB 381 HisLeuAlaTygAspLysAspArgTyrlThrGlnGlnGlnGlnGlnGlnGlnGlnGln 400
QY 1738 CGTTATACAGGAGCAGCCTATACCTGATTAACCTTACAGCAAA 1779
DB 401 ArgTyrlThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 414
  
```

RESULT 6

S04168
 streptokinase A precursor - Streptococcus pyogenes (strain NZ131)
 C:Species: Streptococcus pyogenes
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
 C:Accession: S04168
 R:Huang, T.T.; Malik, H.; Ferretti, J.J.
 Mol. Microbiol. 3, 197-205, 1989
 A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri
 A:Reference number: S04168; MUID:89343623; PMID:2668686
 A:Accession: S04168
 A:Molecule type: DNA
 A:Residues: 1-440 <HDA>
 A:Cross-references: UNIPROT:Q57391; UNIPARC:UPI0000BD04A; EMBL:X51517; NID:g47437; PID
 F:27-440/Product: streptokinase A #status predicted <MNT>
 A:Gene: ska
 C:Superfamily: streptokinase
 P:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-440/Product: streptokinase A #status predicted <MNT>

Alignment Scores:

Pred. No.: 4.8e-118 Length: 440
 Score: 1838.00 Matches: 354
 Percent Similarity: 90.82% Conservative: 22
 Best Local Similarity: 85.51% Mismatches: 38
 Query Match: 58.65% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x S04168 (1-440)

```

QY 538 ATTGCTGACCTGAGTGGCTGCTGACCGTTCATCTGTCAACAACAGCAATTGGTTGTT 597
DB 27 TlelaaglyTyrgLuttrpleuAspArgProSerValAsnSerGlnLeuVal 46
QY 598 AGGTTGCTGCTGCTGAGGGAACAACAAGCACTTGTCTTAATTTTGAATC 657
DB 47 SerValaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 66
QY 658 GATCTTAACATCAAGACCTGCTCATGAGAAAGACAGACAGGCTTAAGTCAAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlnGlnGlnGlnGlnGln 86
QY 718 AAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
DB 87 LysProPheAlaThrAspLysGlyAlaMetSerHisLysLeuGlnGlnGlnGlnGlnGln 106
QY 778 AAGGCTATTCAGAAACAATTTGATGCGAGTCAAGTCAAGACGACCTTGAAGTCTG 837
DB 107 LysAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 126
QY 838 ATTGATTTTGAAGCGATGCAACCATTAATGATGAAACGCAAGGCTTACTTGTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlnGlnGlnGlnGlnGlnGln 146
QY 898 AAAGATGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
  
```

Db 147 LysAspAspSerValThrLeuProThrglnProValGlnGluPheLeuSerGlyHis 166
 QY 958 GTGCGCGTATAGCATATTAAGAAAAACAATACAAACCAAGCAAACTGTGTGATG 1017
 Db 167 ValArgValLysProTyrGlnProLysAlaValHisAspSerAlaGluArgValAsnVal 186
 QY 1018 GAATATCTGTATCAGTTTATCTCCCTTAAACCTGATGACGATTTTCAGACAGGCTCAAA 1077
 Db 167 AsnTyrGluValSerPheValSerGluThrglnLysAspLeuAspPheThrProLeuLeuArg 206
 QY 1078 GATCTAAGCTATTGAAAACTAGTATCGGTGACACATCACTCAATCTCAAGATTACTA 1137
 Db 207 AsnGlnTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerSerGlnGluLeuAla 226
 QY 1138 GCTCAAGCAAGACGATTTTAAACAAACCAACCCGCGCTATACGATTTTATGAACCTGAC 1197
 Db 227 AlaIleAlaGlnPheIleLeuSerLysLysAspAspTyrIleIleThrLysAspAsp 246
 QY 1198 TCCTCAATCGTACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
 Db 247 SerSerIleValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGlu 266
 QY 1258 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGGTCTG 1317
 Db 267 PheThrTyrHisIleLysAspArgGlnAlaTyrLysAlaAsnSerLysThrGlyIle 286
 QY 1318 AATGAAGAATAAACAACATGACCTGATCTGTGAAATATTTACGTCCTTAAAAAGG 1377
 Db 287 GlnGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrLysValLeuLysLysGly 306
 QY 1378 GAAAAGCGGTATGATCCCTTTATGCGCAGTCACTTGAACCTGTTCAACATCAATCGTT 1437
 Db 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
 QY 1438 GATGTCGATACCAACGAATGCTAAACAGTACAGCTTAAACAGTACAGCAACCTAAC 1497
 Db 327 AspValAsnThrAsnLysLeuLeuLysSerGlnGlnLeuThrIleAsnGluArgAsn 346
 QY 1498 TTTAAGCTTCAAGATTTATACATCTCTGTGATAGGCTTAACTCTTCAACCAATCTC 1557
 Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1558 GATGCTTTTGTATTTAGGACTATACCTTAACCTGGAAGTAGAGATATATACAGATGAC 1617
 Db 367 AspAlaAspGlyIleLeuAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 386
 QY 1618 ACCAACCGTATCATACCGTTTATATGCGCAAGCGAACCCGAAAGAGAGATGCTACTAT 1677
 Db 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyLysAsnHisAspTyr 406
 QY 1678 CATTAGCCTTATGATTAAGATGTTATACCGAAGAAAGAGAGATTTACGCTACCTG 1737
 Db 407 HisLeuAlaLysAspLysAspArgTyrThrGlnGlnGlnGluValTyrSerTyrLeu 426
 QY 1738 CGTTATACGGGACACCTTATACCTGATTAACCTTAACAGACAAA 1779
 Db 427 ArgTyrThrGlyThrProIleProAspAsnProLysAspLys 440
 RESULT 7
 S53334
 streptokinase - Streptococcus sp.
 C:Species: Streptococcus sp.
 C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 C:Accession: S53334
 R:Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
 Biochem. J. 304, 235-241, 1994
 A:Title: Function of streptokinase fragments in plasminogen activation.
 A:Reference number: S53334; MUID:95091634; PMID:798939
 A:Accession: S53334
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-197 <SH1>
 A:Cross-references: UNIPROT:Q7M115; UNIPARC:UPI000017AC2C

Alignment Scores:
 Pred. No.: 2,656-46 Length: 197
 Score: 788.50 Matches: 193
 Percent Similarity: 47.34% Conservative: 3
 Best Local Similarity: 46.62% Mismatches: 1
 Query Match: 25.16% Indels: 217
 DB: 2 Gaps: 11
 US-09-940-235-11 (1-1782) x S53334 (1-197)
 QY 538 ATTGCTGACCTGATGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATTGGTTGT 597
 Db 1 IleAlaGlyProGluThrPheLeuAspArgProSerValAsn----- 14
 QY 598 AGCGTTGCTGTACTGTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATTC 657
 Db 14----- 14
 QY 658 GATCTAATCATCAGACCTGCTCATGAGAGAAAGAGAGAGGCTTAAGTCAAAATCA 717
 Db 15-----Ser 15
 QY 718 AAACCAATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTACTA 777
 Db 16 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 35
 QY 778 AAGCTATTTCAGACAAATGATCGTCAACGTCCACAGTACACGACGACTACTTGAAGTTC 837
 Db 36 LysAlaIleGlnThrProGlnIleLeu----- 43
 QY 838 ATTGATTTTCCAGACGATGCAACCATTTACTGATGCAACCGCAAGGTCTACTTGTCTGAC 897
 Db 44-----AsnGlyLysValTyrPheAlaAsp 51
 QY 898 AAAGATGTTGCGTTAACCTTGCCGACCCCAACTGTGCCAAGAAATTTTGTAAAGGACAT 957
 Db 52 LysAspGlySerValThr----- 57
 QY 958 GTGCGCGTTTACCATTAATAAAGAAAAACAATAAACAAGAGAAATCTGTGATGTG 1017
 Db 58-----GlnLysProIleGlnAsnGlnAlaLysSerValAspVal 70
 QY 1018 GAATATACGTATACAGTTTACTCCCTTAAACCTGATGACGATTTTGACACAGGCTCAAA 1077
 Db 71 Gln----- 71
 QY 1078 GATACTAAGCTATTGAAAAACATAGCTATCGTGACACATCATCTCAAGATTACTA 1137
 Db 71----- 71
 QY 1138 GCTCAAGACAAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTGAACGTGAC 1197
 Db 72-----AsnHisProGlyTyrThrIleTyrGlnLysAsp 82
 QY 1198 TCCTCAATGCTCACTATGACATGACATTTTCCGTACGATTTTACCAATGATGACAG 1257
 Db 83 SerSerIleVal-----ThrLeuProMetAspGlnGlu 94
 QY 1258 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATTCATTAATAAATCTGGTCTG 1317
 Db 95 PheThrTyr-----SerGlyLeu 100
 QY 1318 AATGAAGAATAAACAACATGACCTGATCTGTAGAAATATTACGTCCTTAAAAAGG 1377
 Db 101 AsnGlnGlnIleAsnThrAspLeuIleSer----- 111
 QY 1378 GAAAAGCGGTATGATCCCTTATGCGAGTCACTTGAACCTGTCACCATCAAAATACGTT 1437
 Db 112-----TyrVal 113
 QY 1438 GATGTCGATCAACGAATTTGTTAAAAAGTACAGCTTTTACAGCTACGACGATAC 1497

Db 114 AspValAspThrAsnGluLeuLeuSerGluLeuThrAlaSerGluArgAsn 133
Qy 1498 TTAGACTTCAGAGATTATATGATCTCTGATAGGCTAACTACTTACAAATCTC 1557
Db 134 -----AspLeuTyraAspProAspArgAspIleAlaIleuLeuTyraAsnAsnLeu 149
Qy 1558 GATGCTTTGGTATTTATGACTATACCTTAACTGAAAAAGTAGAGATATCAGATGAC 1617
Db 150 AspAlaAspGluIleMet----- 155
Qy 1618 ACCAACCGATGCTATACCGTTTATANGGACGACCCGAGAGAGAAATGCTAGCTAT 1677
Db 156 -----IleIleThrValTyraMetGlyValArgProGluGluAsnAla----- 170
Qy 1678 CATTAAGCTATGATTAAGATCGTTATATACGAGAAAGAAAGATTACAGTACTG 1737
Db 171 -----AspArgTyThrGluGluGluGluValTyraSerTyLeu 184
Qy 1738 CGTTATACAGGACACCTATATACCTGATTAACCTTACGACAAA 1779
Db 185 ---TyThrGlyThrProIleProAspAsnProAspAspLeu 197

RESULT 8

S77671
Streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
C:Accession: S77671; S77672
R:Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittem, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77671
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPARC:UP10000000562; EMBL:U25853; NID:g818908; PIDN:AA85729.1; PI A:Experimental source: strain ET1/ML
A:Note: allele 2
A:Accession: S77672
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPARC:UP10000000562; EMBL:U25854; NID:g818910; PIDN:AA85730.1; PI A:Experimental source: strain B2/M3
A:Note: allele 3
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 8.42e-35 Length: 128
Score: 620.00 Matches: 118
Percent Similarity: 95.31% Conservative: 4
Best Local Similarity: 92.19% Mismatches: 6
Query Match: 19.78% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x S77671 (1-128)

Qy 958 GTGGCGCTTGAACCATTAATAAGAAAAACCAATACAAACGAGAAATCTGTTATGTG 1017
Db 1 ValArgValArgProTyrrLySGluysProValGlaAsnGlnAlaIysSerValAspVal 20
Qy 1018 GAATATACCTGTACAGTTTACCTCTTAAACCGATGATGAGATTTCAGACGAGCTCAAA 1077
Db 21 GluTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 40
Qy 1078 GATACTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGATTACTA 1137
Db 41 AsnThrIlySleuLeuTyrrThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60

Qy 1138 GCTCAAGCAACAAGCATTTTAAACAAAAACCAACCGAGCTATACGATTATGACGTGAC 1197
Db 61 AlaGlnAlaGlnSerIleLeuAsnIlyThrIleProGlyTyrrThrIleTyrrGluArgAsp 80
Qy 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTCAGATTATACCAATGATCAAGAG 1257
Db 81 SerSerIleValThrIleAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
Qy 1258 TTTCCTTACCGGTAAATAATCGGGAACAAGCTTTATAGATCAATAAAAAATCTGTCG 1317
Db 101 PheThrTyrrHisValIySAsnArgGluGlnAlaTyrrGluIleAsnProIlyThrGlyIle 120
Qy 1318 AATGAGAATATAAACAACACTGAC 1341
Db 121 LysGluIySthrAsnAsnThrAsp 128

RESULT 9

S77680
Streptokinase A (EC 3.4.-.-) (allele 11) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77680
R:Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittem, T.S.; Sawyer, S.A.; Musser, M.L. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77680
A:Status: translation not shown

A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54687; UNIPARC:UP100000BD2D3; EMBL:U25862; NID:g818926; PI A:Experimental source: strain ET51/ML7
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.: 2.59e-31 Length: 128
Score: 569.00 Matches: 110
Percent Similarity: 91.41% Conservative: 7
Best Local Similarity: 85.94% Mismatches: 11
Query Match: 18.16% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x S77680 (1-128)

Qy 958 GTGGCGCTTGAACCATTAATAAGAAAAACCAATACAAACGAGAAATCTGTTATGTG 1017
Db 1 ValArgValArgProTyrrLySGluysProIleGlnThrProAlaIysSerValAspVal 20
Qy 1018 GAATATACCTGTACAGTTTACCTCTTAAACCGATGATGAGATTTCAGACGAGCTCAAA 1077
Db 21 ArgTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 40
Qy 1078 GATCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGATTACTA 1137
Db 41 AsnThrIlySleuLeuTyrrThrLeuAlaIleGlyIyThrValThrSerGlnGluLeu 60
Qy 1138 GCTCAAGCAACAAGCATTTTAAACAAAAACCAACCGAGCTATACGATTATGACGTGAC 1197
Db 61 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisProAspTyrrThrIleTyrrGluArgAsp 80
Qy 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTCAGATTATACCAATGATCAAGAG 1257
Db 81 SerSerIleValThrIleAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
Qy 1258 TTTCCTTACCGGTAAATAATCGGGAACAAGCTTTATAGATCAATAAAAAATCTGTCG 1317
Db 101 PheThrTyrrHisValIySAsnArgGluGlnAlaTyrrGluIleAsnIlySleuTyrrSerGlyGln 120
Qy 1318 AATGAGAATATAAACAACACTGAC 1341

Db 121 GluGluysThrAsnAsnThrAsp 128

RESULT 10

streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)

C/Species: Streptococcus pyogenes

C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C/Accession: S77688

R/Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.O. Microbiol. 16, 509-519, 1995

A/Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A/Reference number: S77671; PMID:96037795; PMID:756511

A/Accession: S77688

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-128 <KAP>

A/Cross-references: UNIPROT:O54695; UNIPARC:UP100000BDBDC; EMBL:U25870; NID:g818942; PIR

A/Experimental source: strain ET56/M72

C/Genetics:

A/Gene: ska

C/Superfamily: streptokinase

C/Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	1-72e-30	Length:	128
Score:	557.00	Matches:	107
Percent Similarity:	89.84%	Conservative:	8
Best Local Similarity:	83.59%	Mismatches:	13
Query Match:	17.77%	Indels:	0
Gaps:	2	Gaps:	0

US-09-940-235-11 (1-1782) x S77688 (1-128)

Qy 958 GTGCGCGTTAGACCATTAAGAAAAACCAATACAAACCAAGCAATCTGTTGATGTG 1017

Db 1 ValArgValArgProTyrLysGluysProIleGlnThrProAlaLysSerValAspIle 20

Qy 1018 GAATATACCTGTACGTTTACCTCCCTTAACCCGATGACGATTTGACAGGTCTCAAA 1077

Db 21 ArgTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValIleuLys 40

Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137

Db 41 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 60

Qy 1138 GCTCAGACCAAGCAATTTAAACAAAAACACCCAGGCTATACGATTATGAACTGAC 1197

Db 61 AlaGlnAlaGlnSerIleLeuIleGlnSerHisProAspTyrThrIleTyrGluArgAsp 80

Qy 1198 TCCCTCAATCGTCACATGACATGACATTTCCGTCAGATTTTACCAATGATCAAGAG 1257

Db 81 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100

Qy 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATATAAATCTGCTG 1317

Db 101 PheThrTyrHisValAlaLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspIle 120

Qy 1318 AATGAAGAATAAACAACACTGAC 1341

Db 121 LysGluLysThrAsnAsnThrAsp 128

RESULT 11

S77679

streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)

C/Species: Streptococcus pyogenes

C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C/Accession: S77679

R/Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.O. Microbiol. 16, 509-519, 1995

A/Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A/Reference number: S77671; PMID:96037795; PMID:756511

A/Accession: S77679

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-128 <KAP>

A/Cross-references: UNIPROT:O54686; UNIPARC:UP100000BDBDF; EMBL:U25861; NID:g818924; PIR

A/Experimental source: strain ET50/M43

C/Genetics:

A/Gene: ska

C/Superfamily: streptokinase

C/Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	3.77e-30	Length:	128
Score:	552.00	Matches:	106
Percent Similarity:	89.06%	Conservative:	8
Best Local Similarity:	82.81%	Mismatches:	14
Query Match:	17.61%	Indels:	0
Gaps:	2	Gaps:	0

US-09-940-235-11 (1-1782) x S77679 (1-128)

Qy 958 GTGCGCGTTAGACCATTAAGAAAAACCAATACAAACCAAGCAATCTGTTGATGTG 1017

Db 1 ValArgValArgProTyrLysGluysProIleGlnThrProAlaLysSerValAspIle 20

Qy 1018 GAATATACCTGTACGTTTACCTCCCTTAACCCGATGACGATTTGACAGGTCTCAAA 1077

Db 21 ArgTyrAlaValGlnPheThrProLeuAsnProAspAspAspPheThrProValIleuLys 40

Qy 1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1137

Db 41 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 60

Qy 1138 GCTCAGACCAAGCAATTTAAACAAAAACACCCAGGCTATACGATTATGAACTGAC 1197

Db 61 AlaGlnAlaGlnSerIleLeuIleGlnSerHisProAspTyrThrIleTyrGluArgAsp 80

Qy 1198 TCCCTCAATCGTCACATGACATGACATTTCCGTCAGATTTTACCAATGATCAAGAG 1257

Db 81 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 100

Qy 1258 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATATAAATCTGCTG 1317

Db 101 PheThrTyrHisValAlaLysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 120

Qy 1318 AATGAAGAATAAACAACACTGAC 1341

Db 121 LysGluLysThrAsnAsnThrAsp 128

RESULT 12

FNH1

fibronectin precursor (validated) - human

N/Alternate names: fibronectin splice form ED-A

C/Species: Homo sapiens (man)

C/Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2004

C/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R/Dean, D.C.; Bowler, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A/Reference number: A26460; PMID:87115578; PMID:3031656

A/Accession: A26460

A/Molecule type: DNA

A/Residues: 1-49 <DEA>

A/Cross-references: UNIPROT:P02751; UNIPROT:Q14327; UNIPARC:UP1000016A926; GB:M15801; NID

R/Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A/Title: Evolution of the fibronectin gene.

A/Reference number: A26284; PMID:86111901; PMID:3003095

A/Accession: A26284

A/Molecule type: DNA

A/Residues: 1447-1540 <OLD>

A/Cross-references: UNIPARC:UP1000112E37; GB:M12549; NID:g182668

A/Note: the authors translated the codon TTC for residue 1494 as Glu

R/Poellia, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A>Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
A:Reference number: S00848; PMID:88233940; PMID:3375063
A:Accession: S03917
A:Molecule type: DNA
A:Residues: 1594-1767, 'V', 1769-1783 <PAO>
A:Cross-references: UNIPARC:UPI000017432B; EMBL:X07718; NID:g31402
A>Note: the authors translated the codon AAC for residue 1631 as Asp
R.Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A>Title: Donor and acceptor splice signals within an exon of the human fibronectin gene;
A:Reference number: A24854; PMID:87030929; PMID:3770201
A:Accession: A24854
A:Molecule type: DNA
A:Residues: 1992-2147 <VIB>
A:Cross-references: UNIPARC:UPI000017432C; GB:X04530; NID:g31436
R.Gutman, A.; Yamada, K.M.; Kornblith, A.
FEBS Lett. 207, 145-148, 1986
A>Title: Human fibronectin is synthesized as a pre-propolypeptide.
A:Reference number: A24476; PMID:87030890; PMID:3770189
A:Accession: A24476
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-14, 'Q', 16-38 <GUT>
A:Cross-references: UNIPARC:UPI000017432D
R.Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A>Title: Primary structure of human fibronectin: differential splicing may generate at l
A:Reference number: A91008; PMID:85284965; PMID:2992939
A:Accession: A91008
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 32-1344, 1346-2080, 2112-2386 <KOR>
A:Cross-references: UNIPARC:UPI000017432E; UNIPARC:UPI000017432F; GB:X02761
R.Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A>Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
A:Reference number: A93529; PMID:84272258; PMID:6462919
A:Accession: A93529
A:Molecule type: mRNA
A:Residues: 973-2080, 2112-2386 <KO2>
A:Cross-references: UNIPARC:UPI000017432F; UNIPARC:UPI0000174330; GB:X00739
R.Olberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A>Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
A:Reference number: A21011; PMID:83280929; PMID:6688418
A:Accession: A21011
A:Molecule type: mRNA
A:Residues: 1434-1537 <OL2>
A:Cross-references: UNIPARC:UPI0000174331; GB:X00055; NID:g182680; PIDN:AAAS2459.1; PID:
R.Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A>Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with re
A:Reference number: A90495; PMID:85280409; PMID:2992573
A:Accession: A90495
A:Molecule type: mRNA
A:Residues: 1594-2386 <BER>
A:Cross-references: UNIPARC:UPI000016A928; GB:M10905; NID:g182696; PIDN:AAAS2462.1; PID:
R.Umezawa, K.; Kornblith, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A>Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A:Reference number: A22245; PMID:85231203; PMID:2989004
A:Accession: A22245
A:Molecule type: mRNA
A:Residues: 1948-2067 <UME>
A:Cross-references: UNIPARC:UPI000000046A; GB:M27589; NID:g182705; PIDN:AAAS2465.1; PID:
A:Accession: B22245
A:Molecule type: mRNA
A:Residues: 1975-1991, 2017-2039 <UM2>
A:Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
R.Sekiuchi, K.; Kio, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A>Title: Human liver fibronectin complementary DNAs: identification of two different mes
A:Reference number: I52394; PMID:87026578; PMID:3021206

A:Accession: I65273
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SRK>
A:Cross-references: UNIPARC:UPI00000604C; GB:M14060; NID:g182701; PIDN:AAAS2464.1; PID:
R.Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A>Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A:Reference number: A21165; PMID:83221567; PMID:6304699
A:Accession: A21165
A:Molecule type: mRNA
A:Residues: 2291-2386 <KO3>
A:Cross-references: UNIPARC:UPI0000174334; GB:X00799; NID:g182681; PIDN:AAAS2460.1; PID:
R.Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A>Title: Primary structure of human plasma fibronectin.
A:Reference number: A92398; PMID:84032463; PMID:6630202
A:Accession: A92398
A:Molecule type: protein
A:Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
A:Cross-references: UNIPARC:UPI0000174335
R.Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A>Title: Further characterization of the binding of fibronectin to gelatin reveals the
A:Reference number: S34791; PMID:93312001; PMID:8323285
A:Accession: S34791
A:Molecule type: protein
A:Residues: 291-300, 551-560 <GAR2>
A:Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
R.Giffith, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A>Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A:Reference number: A60904; PMID:87019725; PMID:3532418
A:Accession: A60904
A:Molecule type: protein
A:Residues: 293-301 <GRI>
A:Cross-references: UNIPARC:UPI0000174338
R.Calaycay, J.; Pande, H.; Lee, T.; Borsl, L.; Sirl, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A>Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p
A:Reference number: A23901; PMID:8608277; PMID:3900070
A:Accession: A23901
A:Molecule type: protein
A:Residues: 616-677, 'Q', 679-703, 'PR', <CAL>
A:Cross-references: UNIPARC:UPI0000174339
R.Pierichbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A>Title: The cell attachment domain of fibronectin. Determination of the primary structu
A:Reference number: A92386; PMID:82265604; PMID:7050098
A:Accession: A92386
A:Molecule type: protein
A:Residues: 1441-1548 <PR>
A:Cross-references: UNIPARC:UPI0000141CDS
A>Note: residues 1524-1527 are responsible for the cell-binding activity
R.Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A>Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa do
A:Reference number: A32517; PMID:87241275; PMID:3593230
A:Accession: A32517
A:Molecule type: protein
A:Residues: 1589-1630, 'T', 1722-2058 <GAR3>
A:Cross-references: UNIPARC:UPI000017433A
R.Tresselt, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pan
Biochem. J. 214, 731-738, 1991
A>Title: Human plasma fibronectin. Demonstration of structural differences between the
A:Reference number: S14357; PMID:91190085; PMID:2012601
A:Accession: S14357
A:Molecule type: protein
A:Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>
A:Cross-references: UNIPARC:UPI000017433B
R.Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A>Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-da

A:Reference number: A23891; MUID:85261459; PMID:4019516
A:Accession: A23891
A:Molecule type: protein
A:Residues: 2071-2080;2112-2356 <GAR4>
A:Cross-references: UNIPARC:UPI000017433C; UNIPARC:UPI000017433D
C:Comment: The extra domain and connecting strand 3 are subject to developmental and rib
C:Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
action, and transformation.
C:Genetics:
A:Gene: GDB:FNI
A:Cross-references: GDB:119135; OMIM:135600
A:Map position: 2q34-q34
A:Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C:Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology;
C:Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicated
F:127-31/Domin: signal sequence #status predicted <SIG>
F:132-2386/Product: fibronectin #status experimental <MAT>
F:52-87/Domin: fibronectin and heparin binding <FHB>
F:52-87/Domin: fibronectin type I repeat homology <1F1>
F:97-135/Domin: fibronectin type I repeat homology <1F2>
F:141-179/Domin: fibronectin type I repeat homology <1F3>
F:186-225/Domin: fibronectin type I repeat homology <1F4>
F:231-270/Domin: fibronectin type I repeat homology <1F5>
F:308-608/Domin: collagen binding <CBR>
F:308-342/Domin: fibronectin type II repeat homology <1F6>
F:360-401/Domin: fibronectin type II repeat homology <2F1>
F:420-461/Domin: fibronectin type II repeat homology <2F2>
F:470-508/Domin: fibronectin type I repeat homology <1F7>
F:518-555/Domin: fibronectin type I repeat homology <1F8>
F:561-599/Domin: fibronectin type I repeat homology <1F9>

Alignment Scores:
Pred. No.: 1,2e-26 Length: 2386
Score: 503.00 Matches: 101
Percent Similarity: 66.25% Conservative: 5
Best Local Similarity: 63.12% Mismatches: 10
Query Match: 16.05% Indels: 45
DB: 1 Gaps: 3

US-09-940-235-11 (1-1782) x FNHD (1-2386)
QY 254 CCCATAGCTGAGAGAGGTTTGTATCATGCTGCGGAGCTTCCATGTGTCGAGAGAAAG 313
Db 181 ProllealAGLWVSCyPheNaphiSAlaAlaGlyThrSeryValAlaGlyLutrr 200
QY 314 TGGGA-----GAAGCAGC 327
Db 201 TrpGluVysProTyrgInglyTrpMetMetValaapCyThrCysleuGlygluGlyser 220
QY 328 GGAGGCATCATCTTGCACTTCTGAAATAGATGCAACGATGAGACACAGACATCTTAT 387
Db 221 GlyArgIlethrCyThrSerrArgNasAgyCSasAspGlnAspThrArgThrSeryr 240
QY 388 AGAATYGGAGACACTGGAGCAAGAGATATTCAGAAACCTGCTCCAGGCACTTCG 447
Db 241 ArgIleGlyAspThrTrpSerrYsAspAsnAAGlyAsnleuGlnCysIleCyS 260
QY 448 ACAGGCAACGCGCCGAGAGAGTGAAGTGTGAGAGCAGCACTCTGTGAGACACATCG 507
Db 261 ThrGlyangnglyArgGlygluTrpYsGlyuAghIsthrSeryValInthrThrser 280
QY 508 AGCGGATTTGGCCCTTCAACCGATGTTGCTATGCT----- 543
Db 281 SerGlyserGlyProPhethrAspValArgAlaAlaValIrgInProGlnProHisPro 300
QY 544 -----GGACT 549
Db 301 GlnProProProTyrgIyHisCyValInhrAspSergIyValIValTySeryValGlyMet 320
QY 550 GAGTGTGCTAGACCGTTCATCTGTCAACAGCAGCATTTGTTGAGCGTTGTGT 609
Db 321 GlnTrpLeu-----LysThrGlnGlyAnlySglnMetLeuCyThrCySleuGly 337

RESULT 13
FNBO
fibronectin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A26452; B21165; A23292
R:Skortengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.B.; Magnusson, S.
Eur. J. Biochem. 161, 441-453, 1986
A:Title: Complete primary structure of bovine plasma fibronectin.
A:Reference number: A26452; MUID:87054047; PMID:3780752
A:Accession: A26452
A:Molecule type: protein
A:Residues: 1-2265 <SKO>
A:Cross-references: UNIPROT:P07589; UNIPARC:UPI000012A7BE
R:Kornilint, A.R.; Vide-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A:Reference number: A21165; MUID:83221567; PMID:6304699
A:Accession: B21165
A:Molecule type: mRNA
A:Residues: 2170-2265 <KOR>
A:Cross-references: UNIPARC:UPI000016C30C; GB:K00800; NID:G163055; PIDN:AAA30521.2; PID:
R:Pettersen, T.B.; Thogersen, H.C.; Skortengaard, K.; Vide-Pedersen, K.; Sahl, P.; Sottr
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A:Title: Partial primary structure of bovine plasma fibronectin: three types of interna
A:Reference number: A23292; MUID:83117805; PMID:6218503
A:Accession: A23292
A:Molecule type: protein
A:Residues: 1-16, 'C', 18-20, 'S', 22-43;447-463;1367-1517;1567-1673;2062-2176, 'N', 2178-2265
F:21-241/Domin: fibrin and heparin binding <FBR>
F:21-56/Domin: fibronectin type I repeat homology <1F1>
F:66-104/Domin: fibronectin type I repeat homology <1F2>
F:110-148/Domin: fibronectin type I repeat homology <1F3>
F:155-194/Domin: fibronectin type I repeat homology <1F4>
F:200-239/Domin: fibronectin type I repeat homology <1F5>
F:277-577/Domin: collagen binding <CBR>
F:389-430/Domin: fibronectin type I repeat homology <1F6>
F:329-370/Domin: fibronectin type II repeat homology <2F2>
F:439-477/Domin: fibronectin type II repeat homology <1F7>
F:487-524/Domin: fibronectin type I repeat homology <1F8>
F:530-568/Domin: fibronectin type I repeat homology <1F9>
F:578-661/Domin: fibronectin type II repeat homology <FN3A>
F:688-770/Domin: fibronectin type III repeat homology <FN3B>
F:779-860/Domin: fibronectin type III repeat homology <FN3C>
F:875-957/Domin: fibronectin type III repeat homology <FN3D>
F:965-1046/Domin: fibronectin type III repeat homology <FN3E>
F:1055-1134/Domin: fibronectin type III repeat homology <FN3F>
F:1142-1227/Domin: fibronectin type III repeat homology <FN3G>
F:1235-1318/Domin: fibronectin type III repeat homology <FN3H>
F:1326-1404/Domin: fibronectin type III repeat homology <GN3I>
F:1410-1517/Domin: cell attachment <CAD>
F:1416-1502/Domin: fibronectin type III repeat homology <FN3J>
F:1493-1495/Domin: cell attachment (R-G-D) motif
F:1510-1592/Domin: fibronectin type III repeat homology <FN3K>
F:1600-1870/Domin: heparin binding <HB2>
F:1600-1682/Domin: fibronectin type III repeat homology <FN3L>
F:1692-1773/Domin: fibronectin type III repeat homology <FN3M>
F:1781-1863/Domin: fibronectin type III repeat homology <FN3N>
F:1970-1972/Domin: cell attachment (R-G-D) motif
F:1982-2062/Domin: fibronectin type III repeat homology <FN3O>
F:1985-2226/Domin: fibronectin type III repeat homology <PB2>
F:2085-2124/Domin: fibronectin type I repeat homology <1F10>
F:2130-2167/Domin: fibronectin type I repeat homology <1F11>

F1537-1619/Domain: fibronectin type III repeat homology <FN3X>
 F1614-1616/Region: cell attachment (R-G-D) motif
 F1631-1713/Domain: fibronectin type III repeat homology <FN3L>
 F1721-1803/Domain: fibronectin type III repeat homology <FN3M>
 F1811-1893/Domain: fibronectin type III repeat homology <FN3O>
 F1903-1984/Domain: fibronectin type III repeat homology <FN3O>
 F1992-2074/Domain: fibronectin type III repeat homology <FN3O>
 F2181-2183/Region: cell attachment (R-G-D) motif
 F2193-2273/Domain: fibronectin type III repeat homology <FN3O>
 F2226-2235/Domain: fibronectin type I repeat homology <FI10>
 F2341-2378/Domain: fibronectin type I repeat homology <FI11>
 F2385-2420/Domain: fibronectin type I repeat homology <FI12>
 F253-79, 77-68, 98-124, 136, 142-170, 168-180, 187-216, 214-226, 232-261, 259-271, 308-335, 333
 F2458/Disulfide bonds: 2458-2411, 2409-2420/Disulfide bonds: #status predicted
 F2462/Disulfide bonds: interchain (to 2458) #status predicted

Alignment Scores:

Pred. No.:	2, 5e-23	Length:	2477
Score:	454.50	Matches:	92
Percent Similarity:	63.75%	Conservative:	10
Best Local Similarity:	57.50%	Mismatches:	13
Query Match:	14.50%	Indels:	46
DB:	2	Gaps:	4

US-09-940-235-11 (1-1782) x S14428 (1-2477)

```

OY 254 CCCATAGCTGAGAGAGTGTTCATCATCTGCTGGAGCTTCTTANGTGTGGAGAAACG 313
Db 182 ProilealagluysCyshpshsalaalaglyThrseryValValGlyIutThr 201
OY 314 TGGGA-----GAGGCAGC 327
Db 202 TrpGluLysProTyrglnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlyAam 221
OY 328 GAGAGCATCACTTGGACTTCTTAAATAGATGCAACATGACAGACACAGACATCTTAT 387
Db 222 GlyAglIethrCysThrSerArgshsahgCysAsnAspGlnhspThrArgThrSerTy 241
OY 388 AGAATGAGACACCTGAGAGACAGATTAATCGAGAAACCTGCTCGATGACATCTGC 447
Db 242 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysValCys 261
OY 448 ACAAGCAACGCCCGAGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACACATCG 507
Db 262 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisVal---LeuGlnSerAlaSer 280
OY 508 AGCGGATCGGCCCTTCACCGATGTTGCTATTGCT----- 543
Db 281 AlaGlySerGlySerPheThrAspValAlaArgThrAlaIleTyrglnProGlnThrHisPro 300
OY 544 -----GGACCT 549
Db 301 GlnProAlaProTyrglyHisCysValThrAspSerGlyValValTyrservalGlyMet 320
OY 550 GAGTGGCTGCTAGACCGTTCATCTGCAACAACAGCCAAATTGGTGTAGCGTGTGCTGT 609
Db 321 GlnTrpLeu-----LysSerGlnGlyAspLysGlnMetLeuGlySerThrCysLeuGly 337

```

RESULT 15

S77676

streptokinase A (EC 3.4.-.-) (allele 7) - Streptococcus pyogenes (fragment)

C1Species: Streptococcus pyogenes

C1Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C1Accession: S77676

R1Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.

Mol. Microbiol. 16, 509-519, 1995

A1Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus

A1Reference number: S77671; PMID:96037795; PMID:7565111

A1Status: translation not shown

A1Molecule type: DNA

A1Residues: 1-128 <KAP>

A1Cross-references: UNIPROT:O54683; UNIPARC:UPI0000087143; EMBL:U25858; NID:G818918; PDB:
 A1Experimental source: Etl6/M66
 C1Genetics:
 A1Gene: ska
 C1Superfamily: streptokinase
 C1Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:

Pred. No.:	2, 51e-21	Length:	128
Score:	423.00	Matches:	85
Percent Similarity:	75.78%	Conservative:	12
Best Local Similarity:	66.41%	Mismatches:	31
Query Match:	13.50%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-11 (1-1782) x S77676 (1-128)

```

OY 958 GTGCGGCTTGAAGCATTAATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1017
Db 1 ValArgValArgProTyrglnProLysAlaValHisAsnSerAlaGluArgValAsnVal 20
OY 1018 GAATTAATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGAGACAGGCTTCAA 1077
Db 21 AsnTyrgluValSerPheValSerGluThrGlyAspLeuAspPheThrProSerLeuArg 40
OY 1078 GATTAATAGCTATTGAAAACACTAGCTATCGTGACACCACTCAATCTCAAGAAATTA 1137
Db 41 AspArgTyrglnHisLeuThrThrLeuAlaValGlyAspSerLeuSerGlnGluLeuAla 60
OY 1138 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCGAGGCTATACGATTTATGAACGTGAC 1197
Db 61 AlaIleAlaGlnPheIleLeuSerLysGluHisProAspTyrlleIleThrLysArgAsp 80
OY 1198 TCCTCAATGCTCACTCATGACAAATGACATTTTCGATGACATTTTACCAATGATCAAG 1257
Db 81 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 100
OY 1258 TTACTTACCGGTGTTAAAAATCGGAAACAGCTTATAGATCAATAAAAATCGTCTG 1317
Db 101 PheThrTyrglnHisValLysAsnArgGluGlnAlaTyrglyLeuAsnLysSerGlyGln 120
OY 1318 AATGAAGAAATAAACCAACTGAC 1341
Db 121 LysGluLysIleAsnAsnThrAsp 128

```

Search completed: January 28, 2006, 02:37:54
 Job time : 44.3089 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 192.043 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-11

Perfect score: 3134
Sequence: 1 tcgcttcacgctcgtcgcg.....ataacctaacgacaataa 1782

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p model -DEV=rlp
Q=/sgn2.1/USPTO_epool_p/US09940235/runat_27012006_144218_27563/app_query.fasta_1.7708
-DB=UniProt -QFMT=faстан -SUFIX=rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIG=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CGN_1.1.1355 @runat_27012006_144218_27563 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158	68.9	440	1 STRP_STRQ	P00779 streptococc
2	2129	67.9	414	2 Q53284_STRQ	Q53284 streptococc
3	2117	67.5	436	2 Q6UK57_STRQ	Q6UK57 streptococc
4	2116	67.5	440	1 STRP_STRS1	P10519 streptococc
5	1949	62.2	440	2 Q7X0Y1_STRPY	Q7X0Y1 streptococc
6	1937	61.8	440	1 STRP_STRPY	P10520 streptococc
7	1937	61.8	440	2 Q532X6_STRPY	Q532X6 streptococc
8	1937	61.8	440	2 Q8K5R8_STRP3	Q8K5R8 streptococc
9	1930	61.6	440	2 Q7X0Y2_STRPY	Q7X0Y2 streptococc
10	1929	61.6	440	2 Q7X0Y8_STRPY	Q7X0Y8 streptococc
11	1926	61.5	440	2 Q7X0Y7_STRPY	Q7X0Y7 streptococc
12	1926	61.5	440	2 Q5X9T6_STRP6	Q5X9T6 streptococc
13	1923	61.4	440	2 Q8N2A6_STRP8	Q8N2A6 streptococc
14	1880	60.0	440	2 Q7X0Y3_STRPY	Q7X0Y3 streptococc
15	1872	59.7	440	2 Q7X0Y0_STRPY	Q7X0Y0 streptococc
16	1866	59.5	440	2 Q7X0Y5_STRPY	Q7X0Y5 streptococc

17	1865	59.5	440	2 Q7X0X8_STRPY	Q7X0X8 streptococc
18	1847	58.9	440	2 Q7X0X9_STRPY	Q7X0X9 streptococc
19	1845	58.9	440	2 Q7X0X7_STRPY	Q7X0X7 streptococc
20	1838	58.6	440	2 Q57391_STRPY	Q57391 streptococc
21	1825	58.2	440	2 Q7X0Y6_STRPY	Q7X0Y6 streptococc
22	1810	57.8	440	1 STRQ_STRPY	P96471 streptococc
23	1792	57.2	432	2 Q7X0Y4_STRPY	Q7X0Y4 streptococc
24	788.5	25.2	197	2 Q7M115_STRPY	Q7M115 streptococc
25	728	23.2	141	2 Q7X0X2_STRQ	Q7X0X2 streptococc
26	727	23.2	141	2 Q7X0X3_STRQ	Q7X0X3 streptococc
27	724	23.1	141	2 Q7X0W1_STRQ	Q7X0W1 streptococc
28	723	23.1	141	2 Q7WS87_STRQ	Q7WS87 streptococc
29	717	22.9	141	2 Q7X0W3_STRQ	Q7X0W3 streptococc
30	717	22.9	141	2 Q7X0X4_STRQ	Q7X0X4 streptococc
31	715	22.8	141	2 Q7X0W5_STRQ	Q7X0W5 streptococc
32	714	22.8	141	2 Q7X0X5_STRQ	Q7X0X5 streptococc
33	712	22.7	141	2 Q7X0X6_STRQ	Q7X0X6 streptococc
34	711	22.7	141	2 Q7X0W2_STRQ	Q7X0W2 streptococc
35	711	22.7	141	2 Q7X0W9_STRQ	Q7X0W9 streptococc
36	710	22.7	141	2 Q7X0W6_STRQ	Q7X0W6 streptococc
37	707	22.6	141	2 Q7X0W0_STRQ	Q7X0W0 streptococc
38	705	22.5	141	2 Q7X0V5_STRPY	Q7X0V5 streptococc
39	704	22.5	141	2 Q7X0X0_STRQ	Q7X0X0 streptococc
40	701	22.4	141	2 Q7X0W7_STRQ	Q7X0W7 streptococc
41	697	22.2	141	2 Q7X0W4_STRQ	Q7X0W4 streptococc
42	696	22.2	141	2 Q7X0W8_STRQ	Q7X0W8 streptococc
43	693	22.1	141	2 Q7X0X1_STRQ	Q7X0X1 streptococc
44	690	22.0	141	2 Q7X0R2_STRPY	Q7X0R2 streptococc
45	686	21.9	141	2 Q7X0S4_STRPY	Q7X0S4 streptococc

ALIGNMENTS

RESULT 1
STRP_STRQ STANDARD; PRT, 440 AA.
AC P00779;
ID STRP_STRQ
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Streptococcus C precursor.
GN Name=ekc;
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=85232082; PubMed=2989113; DOI=10.1016/0378-1119(85)90145-3;
RA Maiké H., Roe B.A., Ferretti J.U.;
RT "Nucleotide sequence of the streptokinase gene from Streptococcus
equisimilis H46A.".
RL Gene 34:357-362(1985).
RN [2]
RP PROTEIN SEQUENCE OF 27-440.
RX MEDLINE=83127125; PubMed=6760891;
RA Jackson K.W., Tang J.;
RT "Complete amino acid sequence of streptokinase and its homology with
serine proteases.".
RL Biochemistry 21:6620-6625(1982).
-!- FUNCTION: This protein is not a protease, but it activates
plasminogen by complexing with it. As a potential virulence
factor, it is thought to prevent the formation of effective fibrin
barriers around the site of infection, thereby contributing to the
invasiveness of the cells.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR EMBL; K02986; AAA26974.1; -; Genomic DNA.
DR EMBL; X72832; CAA51351.1; -; Genomic DNA.
DR PIR; A00967; BZSO.
DR PIR; A22801; A22801.
DR PDB; 1BML; X-ray; C/D=38-399.
DR PDB; 1L4D; X-ray; B=40-173.
DR PDB; 1L4Z; X-ray; B=27-173.
DR PDB; 1L0R; X-ray; A/B/C/D=177-314.
DR SMR; P00779; 38-398.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW 3D-structure; Direct protein sequencing; Plasminogen activation;
KW Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase C.
FT VARIANT 195 195 L -> D.
FT VARIANT 207 207 D -> L.
FT CONFLICT 298 300 EKY -> LEYK (in Ref. 2).
FT CONFLICT 438 438 N -> D (in Ref. 2).
FT STRAND 180 180
FT STRAND 184 180
FT TURN 199 200
FT TURN 203 204
FT STRAND 205 205
FT STRAND 209 214
FT TURN 216 217
FT STRAND 219 221
FT HELIX 222 236
FT TURN 238 239
FT STRAND 240 252
FT TURN 253 254
FT STRAND 259 260
FT STRAND 268 270
FT STRAND 275 275
FT STRAND 278 280
FT TURN 282 284
FT STRAND 287 289
FT STRAND 292 304
FT TURN 305 306
SQ SEQUENCE 440 AA; 50140 MW; 8FC1F22648AC77A CRC64;

Alignment Scores:
Pred. No.: 7.63e-142 Length: 440
Score: 2158.00 Matches: 414
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.86% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x STRP_STREQ (1-440)
QY 538 ATTGCTGACCTGATGGCTGCTAGACCGTCCATCTGTCAACAACGCAATTTGTT 597
DB 27 Ileaaglyprrglttrpleuunspatgprosevalashmbnsersglneuvai 46
QY 598 AGCGTGTGCTACTGTTGAGGGAAGATCAAGATTAGTCTTAATTTTGAATC 657
DB 47 SerValaIaaglyThrValgluglyThraAnglnAspIleSerleuylsPhegluile 66
QY 658 GATCTAACATCCAGACCTGCTCATGAGGAAAGACAGACAGAGGCTTAAGTCCAAATCA 717
DB 67 AspIeuThrSerArgProIahIsaglyGlysthrGluglnGlyleuSerProlySer 86
QY 718 AAACCATTTGCTACGATGATGAGCGGATGTCACATTAACCTGAGAAAGCTGACTACTA 777
DB 87 LysProPhealathrasPserGlyalawetSerHslyLeuGluysalAspIeu 106
QY 778 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACGACGACTTATTTGAGTTC 837
DB 107 LysalalIeGlnGlnGlnleuIlealAaenValHisSerAsnAspAspTyrPhegluVal 126

QY 838 ATTGATTTTTCAGACGATGCAACCATTAATGATGCAAAACGCAAGTCTACTTTGCTGAC 897
DB 127 IleaSPhealaserAspIalathrlIethAspArgasnGlyysValTyrPhealAsp 146
QY 898 AAAGATGTTCCGTTAACTTTGCCGACCCAACTGTTCCAGAAATTTTGTAAACGCAAT 957
DB 147 LysAspGlySerValThrlleuProthnglnProValGlnGlnPheleuSerGlyHis 166
QY 958 GTCGCGCTTGACCATTAAGAAACCAATCAAAACCAAGCAAGAAATCTGTGAAGTG 1017
DB 167 ValArgValArgProTyrLysGlnysProIleGlnasnGlnalalySerValAspVal 186
QY 1018 GAATTAATCTAGACTTAATCTCCCTTAACCCCTGATGACGATTTCAACAGGCTCAAA 1077
DB 187 GluTyrThrValGlnPheThrProIleuAsnProAspAspAspPheArgProGlyLeuLys 206
QY 1078 GATTAATAGCTATTGAAAACATAGCTATGCGTGACACATCAATCTCAAGAAATTA 1137
DB 207 AspThrLysleuLeuysfThrlleuAlaIleGlyAspThrIleThrserGlnGluLeu 226
QY 1138 GCTCAAGCAGAAAGCAATTTTAAACAAAACACCCAGGCTATACGATTTTGAACGTGAC 1197
DB 227 AlaGlnAlaGlnserIleleuasnlyAsnHlsProGlyTyrThrIleTyrGluArgAsp 246
QY 1198 TCCCTCAATGCTCACTGATGACATGATCATTTCCGTCGATTTTACCAATGATCAAGAG 1257
DB 247 SerSerIleValThrHlsAspAsnAspIlePheArgThrIleuProMetAspGlnGlu 266
QY 1258 TTTACTTACCGTGTAAATTCGGGAACAGCTTAATGATCAATTAATAATCGTCTCG 1317
DB 267 PheThrTyrArgValIlysasnArgGlnGlnalalyArgIleasnlylsySerGlyLeu 286
QY 1318 AATGAAGAATTAACAACACTGACCTGATCTCGAAGAAATTTACGTCCTTAATAAAGG 1377
DB 287 AsnGluGlnIleasnAsnThrAspIleuIleSerGlnlyserTyrValIleuylslyGly 306
QY 1378 GAAAGCCGATGATCCCTTGTATGCCAGTCACGTGAACGTTCAACATCAATACGTT 1437
DB 307 GluLysProTyrAspProPheAspArgSerHlsleuLysleuPheThrIleLysTyrVal 326
QY 1438 GATGTCGATCAACAAGATTTGCTAAAAAGTACAGACTCTTAACAGCTACGACGATAC 1497
DB 327 AspValAspThrAsnGlnleuLeuLysSerGlnGlnleuThrIalaserGluArgAsn 346
QY 1498 TTAGACTTCAGAGATTTTATTCGATCTCGTATGAGGCTTAATCTCTTCAACATCTC 1557
DB 347 LeuAspPheArgAspIeuTyrAspProArgAspIysalalySerleuLeuTyrAsnAsnLeu 366
QY 1558 GATGCTTTGGTATTTATGACCTATACCTTAACGTGAAGAAATGAGGATTAATCCAGTAC 1617
DB 367 AspAlaPheGlyIleMetAspTyrThrIleThrGlyysValGluAspAsnHlsAspAsp 386
QY 1618 ACCAACCGTATCAATACCGTTTATATGGGCAACGACCCGAGAGGAGAAATGCTAGCTAT 1677
DB 387 ThrAsnArgIleIethrValTyrMetGlyLysArgProIuGlnyuAsnAlaserTyr 406
QY 1678 GATTAGCTTATGATTAAGATGCTTATACCGAAAGAAACGAGAAAGTTTACGCTACTG 1737
DB 407 HisleuAlatyrAspAspArgTyrThnglnGlnGlnArgIuValTyrSerTyrLeu 426
QY 1738 CGTTATACAGGACACCTATACCTGATTAACCTTAAGACAAA 1779
DB 427 ArgTyrThrGlyThrProIleProAspAsnProAsnAspLys 440

RESULT 2
O53284_STREQ PRELIMINARY; PRT; 414 AA.
AC O53284;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SKC-2.
OS Streptococcus equisimilis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OK NCBI_Taxid=119602;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RA Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,
 RA Rubiera R., Pedraza A., Padron G., Antuch W., la Fuente J., et al.;
 RT "High level expression of streptokinase in *Escherichia coli*.";
 RL Biotechnology 0:1138-1142 (1992).
 DR EMBL; S46536; AAC60418.1; -; Genomic_DNA.
 DR PDB; 1CAP; X-ray; A/B/C/D=149-285.
 DR SMR; Q53284; 12-372.
 DR GO; GO:0008243; F:plasmaeionogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 SQ SEQUENCE 414 AA; 47254 MW; F75B5831B766904 CRC64;

Alignment Scores:
 Pred. No.: 7, 96e-140 Length: 414
 Score: 2129.00 Matches: 408
 Percent Similarity: 99.03% Conservative: 2
 Best Local Similarity: 98.55% Mismatches: 4
 Query Match: 67.93% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q53284_STRBQ (1-414)

QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACCAATGGTGT 597
 Db 1 MetAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 558 AGCGTTCCTGCTGCTTTGAGGGGACGATCAAGCACTTACCTTAATTTTGAATC 657
 Db 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPheGluIle 40
 QY 658 GATCTAACATCAACGCTGCTGCTGATGAGAAAGACGACGAAGCTTAATCCAAATCA 717
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnGlnLeuSerProLysSer 60
 QY 718 AAACCATTTGCTACTGATAGTGGCGGATGCTCATTAATTTGAGAAGCTGACTTACTA 777
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeu 80
 QY 778 AAGGCTATTCAAGAACATTTGATGCTTAACGTCACAGTAAACGACACTTCTTGAAGTC 837
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
 QY 838 ATTGATTTTGGCAAGCATGACCACTTACTGATCGAAAGCGCAAGGCTTACTTGGCTGAC 897
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 120
 QY 898 AAAAGATGGTTCGGTAACTTGGCGACCAACCTGCTCCAGAAATTTTGTAAAGCGGACAT 957
 Db 121 LysAspGlySerValThrLeuProInGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 958 GTGCGCGTTAGACCATATTAAGAAAAACAATCAACAACCAAGCGAAATCTGTTGATGTG 1017
 Db 141 ValArgValArgProTyrLysGlyLysPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
 QY 1018 GAATTAATCTGTAAGTTACTCCCTTAAACCTGATGACGATTTGACACCGAGTCTCAAA 1077
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1078 GATTAATGCTATTGAAAACACATAGTATCGGTGACACATCACTCAAGAAATTACTA 1137
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 1138 GCTCAAGCAAGAAGCTTTTAAACAAACCAACCCGAGGCTATCGATTTATGAAAGTGAC 1197
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisPheGlyTyrThrIleTyrGlnLysArg 220

QY 1198 TCCTCAATCGTCACTGACATGACATGACATTTTCCGTACATTTTACCAATGATGACAG 1257
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIle 240
 QY 1258 TTACTTACCGGTGTTAAAAATCGGACACAGCTTATAGATCAATAAAAATCTGGTCTG 1317
 Db 241 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnLysLysSerGlyLeu 260
 QY 1318 AATGAAGAATTAACAACACACGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGG 1377
 Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 280
 QY 1378 GAAAGCGATGATGATCCCTTTGATCGCAGTCACTTGAACCTTCAACCATTAATAGCTT 1437
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1438 GATGTCGATACCAAGATTTGCTAAAAAGTACGACGCTTTTACAGCTAGCGAACTTAC 1497
 Db 301 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 320
 QY 1498 TTGACCTTCAGAGATTTATACGATCTCTGATTAAGCTTAACCTTACTTACACATCTC 1557
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 340
 QY 1558 GATGCTTTGGTATTATGACCTATACCTTAACCTGAAAGTACGAGATTAATCAGATGAC 1617
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlnLysValGlnAspAsnHisAspAsp 360
 QY 1618 ACCAAGCTATCATACCGTTTATATGAGGCAAGGACCCGAAAGAGAAATGCTAGTAT 1677
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyAsnArgProGlnGlnGlnAlaSerTyr 380
 QY 1678 CATTACCTTATGATTAATAAGTGGTATATCCGAAAGAACGAGAGATTTCAGCTTACCTG 1737
 Db 381 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 400
 QY 1738 CGTTATACAGGACACACTTATACCTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1779
 Db 401 ArgTyrThrGlnThrProIleProAspAsnProAsnAspLys 414

RESULT 3
 ID Q6UK57_STRBQ PRELIMINARY; PRT; 436 AA.
 AC Q6UK57;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Streptokinase.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OK NCBI_Taxid=119602;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35666;
 RA Costa C.S., Torres F.A.G., Filho S.A.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY568335; AAQ73571.1; -; Genomic_DNA.
 DR SMR; Q6UK57; 59-394.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasmaeionogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 436 AA; 49658 MW; CA2DD95F957DBF3C CRC64;

Alignment Scores:
 Pred. No.: 5, 5e-139 Length: 436
 Score: 2117.00 Matches: 405

Percent Similarity: 98.55% Conservative: 3
 Best Local Similarity: 97.83% Mismatches: 6
 Query Match: 67.55% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q6UK57_STR0Q (1-436)

```

QY 538 ATTGCTGAGCTGAGTGGCTGAGACCGTCCATCTGTCAACCAAGCAATGGTGT 597
DB 23 IleaaglyProgluThreuleuAspArgProSerValAsnAsnSerGlnLeuVal 42
QY 558 AGCGTGTGCTACTGTTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 657
DB 43 SerValaIaaglyThrValGluGlyThrAsnGlnAspIleSerLeuysPhepegIuile 62
QY 658 GATCTTACATCAAGACCTGCTCAATGAGGAAAGACAGACAGCGCTTAAGTCAAAATCA 717
DB 63 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 82
QY 718 AAACCATTTGCTACTGATAGTGGCGGAGTGTCAATTAACCTTGAGAAAGCTGACTACTA 777
DB 83 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlyValAspLeuLeu 102
QY 778 AAGGCTATTCAAGAACCAATTGATCGCTTAACGTCCACAGTAAACGACACTATTGAGGTC 837
DB 103 LysAlaIaIeGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrrheGlyVal 122
QY 838 ATTGATTTTGCAAGCGATGCAACCATTAATCTGATGAAAGCGGAGGCTTACTTGTGTGAC 897
DB 123 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrrheAlaAsp 142
QY 898 AAAGATGGTGGTGAACCTTGCGCGACCCCAACCTGTCAGAAATTTTGTAAAGCGACAT 957
DB 143 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuysGlyHis 162
QY 958 GTGCGCGTTAGACCATATTAAGAAAAACCAATCAAAACCAAGCGCAATCTGTTGATGTG 1017
DB 163 ValaGValArgProTyrrLysGlyLysProValGlnAsnGlnAlaLysSerValaAspVal 182
QY 1018 GAATTAATCTGTAACGATTACTCCCTTAAACCTGATGACGATTTCAGACCAAGGCTCAAA 1077
DB 183 GlyTyrrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuys 202
QY 1078 GATATAAGCTATTGAAAACTAGTATCGGTGACACCATCAATCTCAAGAAATTAATA 1137
DB 203 AspThrLysLeuLeuysThrLeuAlaIeGlyAspThrIleThrSerGlnGlnLeuLeu 222
QY 1138 GCTCAAGACCAAGACATTTAACAAAAACCAACCCGAGGCTATACGATTATGAACGTGAC 1197
DB 223 AlaGlnAlaGlnSerIleLeuAsnLysThrHisAspGlyTyrrThrIleTyrrGlnAlaArgAsp 242
QY 1198 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1257
DB 243 SerSerIleValIleThrHisAspLysAspIlePheArgThrIleLeuProMetAspGlnGln 262
QY 1258 TTTACTTACCGGTGTTAAAAATCGGAAACAGGCTTATAGGATCAATAAAAATTCGGTCTG 1317
DB 263 PheThrTyrrHisValLysAsnArgGlnGlnAlaTyrrGlnIleAsnLysLysSerGlyLeu 282
QY 1318 AATGAAGAAATTAACCAACGCTGACCTGTGAAATAATATTACGCTTAAAAAAGG 1377
DB 283 AsnGlnGlyIleAsnAsnThrAspLeuIleSerGlnLysTyrrTyrrValLeuLysGly 302
QY 1378 GAAAAGCGGTATGATCCCTTTGATCGACATCTTGAAAATGTTCAACAATACGTT 1437
DB 303 GlyLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrrVal 322
QY 1438 GATGTCGATCCAAAGCAATTCGTTAAAAAGTGAACAGCTCTTAAACAGCTAAGCAAGTAAAC 1497
DB 323 AspValaAsnThrAsnGlnLeuLeuysSerGlnGlnLeuThrAlaSerGlnAlaArgAsn 342
QY 1498 TTAGACTTCAGAGATTTATAGATCCTCGTGAATTAAGGTAACATACTTCAACAACATTC 1557

```

```

DB 343 LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu 362
QY 1558 GATGCTTTTGGTATTATAGCACTATACCTTAACGTGAAAGATGAGGATATACAGATGAC 1617
DB 363 AspAlaPheGlyIleMetAspTyrrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 382
QY 1618 ACCAAGCGTATCAATTAACCGTTTATATGGGACGACCCGAGAGAGAGAAATGCTAGCTAT 1677
DB 383 ThrAsnArgIleIleThrValTyrrMetGlyLysArgProGlnGlyLysAlaSerTyrr 402
QY 1678 CATTAAGCCTATGATTAAGATCGTTATACCGAAGAACAGAGAGATTACAGCTACTG 1737
DB 403 HisLeuAlaTyrrAspLysAspArgTyrrThrGlnGlnGlnAlaGlyValTyrrSerTyrrLeu 422
QY 1738 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAA 1779
DB 423 ArgTyrrThrGlyThrProIleProAspAsnProAsnAspLys 436

```

RESULT 4

```

ID STRP_STRS1 STANDARD; PRT; 440 AA.
AC P10519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Streptokinase G precursor.
OS Name=SkG;
OS Streptococcus sp. (strain 19909).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=69017;
RN [1] NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RP MEDLINE=89160265; PubMed=2922269;
RA Walter F., Siegel M., Maile H.;
RT "Nucleotide sequence of the streptokinase gene from a group-G
  Streptococcus."
RL Nucleic Acids Res. 17:1262-1262(1989).
CC - FUNCTION: This protein is not a protease, but it activates
  plasminogen by complexing with it. As a potential virulence
  factor, it is thought to prevent the formation of effective fibrin
  barriers around the site of infection, thereby contributing to the
  invasiveness of the cells.

```

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

```

CC -----
DR EMBL; X13400; CNA31766.1; -; Genomic_DNA.
DR PIR; S02723; S02723.
DR HSSP; P00779; 1QOR.
DR SMR; P10519; 63-398.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KM Plasminogen activation; Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase G.
SQ SEQUENCE 440 AA; 50199 MW; 5521F8825FE1B6EA CRC64;

```

Alignment Scores:

```

Pred. No.: 6,46e-139      Length: 440
Score: 2116.00      Matches: 406
Percent Similarity: 98.55%      Conservative: 2
Best Local Similarity: 98.07%      Mismatches: 6
Query Match: 67.52%      Indels: 0
DB: 1      Gaps: 0

```

US-09-940-235-11 (1-1782) x STRP_STRS1 (1-440)

```

OY 538 ATTGCTGGAACCTGAGTGGCTGCTAGAACCTGCTCATCTGTCTCAACAACAGCAATGTTGTT 597
DB 27 TlealaglyProgluThrleuLeuAspArgProSerValaAsnAsnSerGlnLeuVala 46
OY 598 AGCGTTGCTGTAACCTGTAAGGGGAGCAATCAGACATTAGTCTTAATTTTGAATC 657
DB 47 SerValaAlaglyThrValaGluGlyThrAsnGlnAspIleSerLeuysPhePheGluIle 66
OY 658 GATCTAAACATCAACGCTGCTCAGAGGAGAAAGACAGAGGAGGCTTAAGTCCAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnLeuSerProLysSer 86
OY 718 AAACATTGCTCATGATAGTGGCGCGATGTCACATTAACCTGAGAAAGCTGACTTACTA 777
DB 87 LysLeuPheAlaThrAspSerGlyAlaMetProHisblyLeuGlnLysAlaAspLeuLeu 106
OY 778 AAGGCTATTCAGAAACAATTTGATCGCTTAAGTCAAGTCAAGACGACTTCTTGAAGTC 837
DB 107 LysAlaIleGlnGlnLeuIleAlaAsnValaHisSerAsnAspArgPheGluVala 126
OY 838 ATTGATTTTGAAGCGACGATCAACCATTAAGTCAAGGAGGCTTCTTGTGTCAG 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValaIlePheAlaAsp 146
OY 898 AAAGATGGTTCGGTAACTTGGCCGACCAACCTGTCAGAAATTTTGTCAAGCGGACAT 957
DB 147 LysAspGlySerValaThrleuProIleGlnProValaGlnLysLeuLeuLysGlyHis 166
OY 958 GTGCGCGTTAGAACATATTAAGAAAAACAATCAAAACAGACGAAATCTGTTGATG 1017
DB 167 ValaGlyValaArgProLysGlyLysProValaGlnAsnGlnAlaLysSerValaAspVala 186
OY 1018 GAATPACTGTCAGTTTACCTCCCTTAACCCGTAAGAGATTTGACACGAGCTCAAA 1077
DB 187 GluLysThrValaGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
OY 1078 GATCTAAAGCTATTTGAAAACTAGCTATCGGTGACACATCACTCTCAAGATTACTA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleHisSerGlnLysLeuLeu 226
OY 1138 GCTCAAGCAACAAGCTTTTAAACAAAAACAACCCAGGCTATACGATTATGAACGTGAC 1197
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyLysThrIleLysGlnLysAsp 246
OY 1198 TCCTCAATGTGACTCATGACATGACATTTTCCGACATTTTACCAATGATGACAG 1257
DB 247 SerSerIleValaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnLys 266
OY 1258 TTACTCTTACCGTGTAAATAATGGGAAACAAGCTTATAGGATCAATTAATAATCGTCTG 1317
DB 267 PheThrLysHisValaLysAsnAsnGlnGlnAlaLysArgLleAsnLysLysSerGlyLeu 286
OY 1318 AATGAAGAATAAACAACAACATGACCTGATCTCTGAGAAATAATTAAGCTCTTAATAAAGG 1377
DB 287 AsnGlnGlnLysAsnAsnThrAspLeuIleSerGlnLysLysLysLysLysLysLys 306
OY 1378 GAAAACCGGTATGATCCCTTGAATGCGAGTCACTGAAAACGTGTCACATCAAAATCGT 1437
DB 307 GluLysProLysAspProPheAspArgSerHisLeuLysLeuPheThrIleLysLysVala 326
OY 1438 GATGTCGATACCAAGAAATTTGCTAAATAAGTGACACTCTTAAACACTAGCAAGCTAAC 1497
DB 327 AspValaAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsn 346
OY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAACTACTTACCAACAATCTC 1557
DB 347 LeuAspPheArgAspLeuLysAspProAlaHisGlyLysThrGlnGlnLysLeuLysLeu 366
OY 1558 GATGCTTTTGGTATTATGACATTAACCTTAAGTGAAGGATGATATATCAAGATGAC 1617
DB 367 AspAlaPheGlyIleMetAspLysLysThrIleuThrGlnLysValaGlnLysAsnHisAspAsp 386
OY 1618 ACCAACCGTATCATTAACGTTTATATATGAGCAAGCGAACCGAAGAGAGAAATGCTAGCTAT 1677

DB 387 ThrAsnArgIleIleThrValaLysMetGlyLysArgProGlnGlnAlaSerLys 406
OY 1678 CATTACCTATGATTAAGAATCGTTATATCCGAAAGAAAGAAAGTTTAACTACTCTG 1737
DB 407 HisLeuAlaLysAspLysAspArgLysThrGlnLysGlnLysValaLysSerLysLeu 426
OY 1738 CGTTATACAGGAGCAACCTATATCACTGATTAACCTTACAGCAAA 1779
DB 427 ArgLysThrGlnLysThrProIleProAspAsnProAsnLys 440

RESULT 5
ID Q7X0Y1_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=eka;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D488;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations." ;
RL J. Bacteriol. 186:110-121(2004)
DR EMBL; AY234137; AAP39957.1; -; genomic_DNA.
DR HSSP; Q53284; IC4P.
DR SMK; Q7X0Y1; 63-398.
DR GO; GO:0016301; F:Kinase activity; IEA.
DR GO; GO:0008243; F:Plasminogen activator activity; IEA.
DR GO; GO:0005515; F:Protein binding; IEA.
DR InterPro; IPR004093; Scaphylokinase.
DR Pfam; PF02821; Staphylokinase.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4BE69647043BAC CRC64;

Alignment Scores:
Pred. No.: 2.83e-127 Length: 440
Score: 1949.00 Matches: 372
Percent Similarity: 94.20% Conservative: 18
Best Local Similarity: 89.86% Mismatches: 24
Query Match: 62.19% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y1_STRPY (1-440)
OY 538 ATTGCTGGAACCTGAGTGGCTGCTAGAACCTGCTCATCTGTCTCAACAACAGCAATGTTGTT 597
DB 27 TlealaglyProgluThrleuLeuAspArgProSerValaAsnAsnSerGlnLeuVala 46
OY 598 AGCGTTGCTGTAACCTGTAAGGGGAGCAATCAGACATTAGTCTTAATTTTGAATC 657
DB 47 SerValaAlaglyThrValaGluGlyThrAsnGlnAspIleSerLeuysPhePheGluIle 66
OY 658 GATCTAAACATCAACGCTGCTCAGAGGAGAAAGACAGAGGAGGCTTAAGTCCAAATCA 717
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnLeuSerProLysSer 86
OY 718 AAACATTGCTCATGATAGTGGCGCGATGTCACATTAACCTGAGAAAGCTGACTTACTA 777
DB 87 LysLeuPheAlaThrAspSerGlyAlaMetProHisblyLeuGlnLysAlaAspLeuLeu 106
OY 838 ATTGATTTTGAAGCGACGATCAACCATTAAGTCAAGGAGGCTTCTTGTGTCAG 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValaIlePheAlaAsp 146
OY 898 AAAGATGGTTCGGTAACTTGGCCGACCAACCTGTCAGAAATTTTGTCAAGCGGACAT 957
DB 147 LysAspGlySerValaThrleuProIleGlnProValaGlnLysLeuLeuLysGlyHis 166
OY 958 GTGCGCGTTAGAACATATTAAGAAAAACAATCAAAACAGACGAAATCTGTTGATG 1017
DB 167 ValaGlyValaArgProLysGlyLysProValaGlnAsnGlnAlaLysSerValaAspVala 186
OY 1018 GAATPACTGTCAGTTTACCTCCCTTAACCCGTAAGAGATTTGACACGAGCTCAAA 1077
DB 187 GluLysThrValaGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
OY 1078 GATCTAAAGCTATTTGAAAACTAGCTATCGGTGACACATCACTCTCAAGATTACTA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleHisSerGlnLysLeuLeu 226
OY 1138 GCTCAAGCAACAAGCTTTTAAACAAAAACAACCCAGGCTATACGATTATGAACGTGAC 1197
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyLysThrIleLysGlnLysAsp 246
OY 1198 TCCTCAATGTGACTCATGACATGACATTTTCCGACATTTTACCAATGATGACAG 1257
DB 247 SerSerIleValaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnLys 266
OY 1258 TTACTCTTACCGTGTAAATAATGGGAAACAAGCTTATAGGATCAATTAATAATCGTCTG 1317
DB 267 PheThrLysHisValaLysAsnAsnGlnGlnAlaLysArgLleAsnLysLysSerGlyLeu 286
OY 1318 AATGAAGAATAAACAACAACATGACCTGATCTCTGAGAAATAATTAAGCTCTTAATAAAGG 1377
DB 287 AsnGlnGlnLysAsnAsnThrAspLeuIleSerGlnLysLysLysLysLysLysLys 306
OY 1378 GAAAACCGGTATGATCCCTTGAATGCGAGTCACTGAAAACGTGTCACATCAAAATCGT 1437
DB 307 GluLysProLysAspProPheAspArgSerHisLeuLysLeuPheThrIleLysLysVala 326
OY 1438 GATGTCGATACCAAGAAATTTGCTAAATAAGTGACACTCTTAAACACTAGCAAGCTAAC 1497
DB 327 AspValaAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsn 346
OY 1498 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAACTACTTACCAACAATCTC 1557
DB 347 LeuAspPheArgAspLeuLysAspProAlaHisGlyLysThrGlnGlnLysLeuLysLeu 366
OY 1558 GATGCTTTTGGTATTATGACATTAACCTTAAGTGAAGGATGATATATCAAGATGAC 1617
DB 367 AspAlaPheGlyIleMetAspLysLysThrIleuThrGlnLysValaGlnLysAsnHisAspAsp 386
OY 1618 ACCAACCGTATCATTAACGTTTATATATGAGCAAGCGAACCGAAGAGAGAAATGCTAGCTAT 1677

```

```

Db      107  LysAlaIleGlnIysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Oy      838  ATTGATTTTGAAGCGATCAACCATTTATGATCGAAACGGGACGGCTTATGCTTGAC 897
Db      127  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyGlnValTyrPheAlaAsp 146
Oy      898  AAAAGATGCTTCGGTAACTTTCGCCAGCCCAACCTGTCCAAAGATTTTTCCTAACCGGACAT 957
Db      147  LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Oy      958  GTGGCGGCTTGAACCATATAAAGAAAAACCATACAAACCGGAAATCTGTTGATGCG 1017
Db      167  ValArgValArgProTyrIysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
Oy      1018  GAAATATACGTGACAGTTTACTCCCTTAAACCCGTATGACGATTTTGACACAGCTTCAAA 1077
Db      187  LysTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
Oy      1078  GATTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCAATCTCAAGATTAATA 1137
Db      207  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
Oy      1138  GCTCAAGCAACAAGCATTTTAAACAAAACCAACCGCTATACGATTTATGAACTGAC 1197
Db      227  AlaGlnAlaGlnSerIleLeuLysThrHisProGlyTyrThrIleTyrGluArgAsp 246
Oy      1198  TCCCTAATCGTCATCAATGACATGATTTCCCGTACGATTTTACCATGATGATCAAG 1257
Db      247  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Oy      1258  TTTACTTACCGTGTAAATAATCGGGAACAGCTTATAGGATCAATAAATACTGCTG 1317
Db      267  PheThrTyrArgValLysAspArgGluGlnAlaTyrGlyIleAsnLysSerGlyLeu 286
Oy      1318  AATGAGAATAATCAACAACATGACCTGATCTGTAGAAAATATTATCGTCTTAAAAAGG 1377
Db      287  AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrIleLeuLysLysGly 306
Oy      1378  GAAAAGCCGATGATCCCTTATGATCGGACATCTGAAAACCTGTCACCATCAATACG 1437
Db      307  GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
Oy      1438  GATGTCGATTAACCAAGATTTGCTTAAAGAGAGAGCTCTTAAACAGCTACGGAAGTAC 1497
Db      327  AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Oy      1498  TTAAGCTTCAGAGATTATACATCTCTCGTGAATAGGCTTAACTACTTACCAATCTTC 1557
Db      347  LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Oy      1558  GATGCTTTTGGTATTATGACATATACCTTAACTGAAAAATGAGAGATTAATCAAGTAC 1617
Db      367  AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluLysAsnHisAspLys 386
Oy      1618  ACCAACCGATATCAACCGTTTATATGAGGACGACCGGACGAGAGAGATGCTACTAT 1677
Db      387  AsnAsnAspGlnLeuAlaThrValIleTyrMetGlyLysArgProLysGlnAlaLysGlySerTyr 406
Oy      1678  CATTAGCCTATGATTAAGATGCTTATACCGAAGAGAAAGAGATTAAGCTTACGTCG 1737
Db      407  HisLeuAlaTyrAspLysAspProTyrThrGlnGluGluArgLysAlaLysSerTyrLeu 426
Oy      1738  CGTTATACAGGAGCACTTATACCTGATTAACCTTAAACGACAAA 1779
Db      427  ArgAspThrGlnThrProIleProAspAsnProLysAspLys 440

```

```

DE      Streptokinase A precursor.
GN      Name=ska; OrderedLocNames=SPY1979;
OS      Streptococcus pyogenes.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus.
OX      NCBI_TaxID=1314;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=SF130/13 / Serotype M1;
RX      MEDLINE=89160264; PubMed=2646590;
RA      Walter F., Siegel M., Walke H.;
RT      "Nucleotide sequence of the streptokinase gene from a Streptococcus
RL      pyogenes type 1 strain."
RL      Nucleic Acids Res. 17:1261-1261(1989).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX      MEDLINE=1192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA      Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA      Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA      Qian Y., Jia H.G., Nejar P.Z., Ren Q., Zhu H., Song L., White J.,
RA      Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT      "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RT      Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC      -1- FUNCTION: This protein is not a protease, but it activates
CC      plasminogen by complexing with it. As a potential virulence
CC      factor, it is thought to prevent the formation of effective fibrin
CC      barriers around the site of infection, thereby contributing to the
CC      invasiveness of the cells.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
DR      EMBL: X13399; CAA31765.1; -; Genomic DNA.
DR      EMBL: AB006620; AAK34665.1; -; Genomic DNA.
DR      PIR: S02724; S02724.
DR      HSSP: Q53284; 1CAP.
DR      SMR: P10520; 63-398.
DR      InterPro: IPR004093; Staphylokinase.
DR      InterPro: IPR008124; Streptokinase.
DR      Pfam: PF02821; Staphylokinase; 3.
DR      PRINTS: PR01753; STREPKINASE.
KW      Complete proteome; Plasminogen activation; Signal; Virulence.
FT      CHAIN 1..26
FT      FT 27..440
FT      CONFLICT 163..163 L -> V (in Ref. 1).
FT      CONFLICT 345..345 R -> G (in Ref. 1).
FT      CONFLICT 373..373 D -> N (in Ref. 1).
FT      CONFLICT 428..428 D -> Y (in Ref. 1).
FT      CONFLICT 438..438 K -> N (in Ref. 1).
SQ      SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;

Alignment Scores:
Pred. No.: 1,94e-126 Length: 440
Score: 1937.00 Matches: 367
Percent Similarity: 93.48% Conservative: 20
Best Local Similarity: 88.65% Mismatches: 27
Query Match: 61.81% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-11 (1-1782) x STRP_STRPY (1-440)
Oy      538  ATTGCTGAGCGCTGAGGCTGCTAGACCGTCCATCTGTCACAAACGCAATGTTGTT 597
Db      27  IleAlaGlyTyrGlyTyrLeuProAspArgProIleAsnAsnSerGlnLeuValVal 46
Oy      598  AGCGTGTCTGACTGTTGATGAGGAGCAATCAACATCAAGATTAAGCTTAATTTTGAATC 657
Db      47  SerMetAlaGlyIleValGluGlyThrAspLysLysValPheIleAsnPhePheGluIle 66

```

```

QY 658 GATCTAATCATCAGCACTGCTCATGAGAAAGACAGACGAGCTTAAGTCCAAATCA 717
DB 67 AspleuThrSerGlnProAlaHisGlyGlySerThrGlnGlnGlyLeuSerProLysSer 86
QY 718 AAACCAATTGCTAGTGTAGTGGGAGATGCATTAACCTTGAAAGCGAATTCTTA 777
DB 87 LysProPheMetThrAspAsnGlyAlaMetProHisLysLeuGlyValAspLeuLeu 106
QY 778 AAGCTATTTCAGAAACAATTGATCGCTAACGTCACAGTAACGACGACTTACTTGAGGTC 837
DB 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 838 ATTGATTTTGCAAGCGATGCAACCATTAAGTGAAGCGGCAAGGTCTTACTTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 146
QY 898 AAAGATGTTGGTGAACCTTGCCGACCACTGTCACAAAGATTTTGGTGAAGGACAT 957
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuLysGlyHis 166
QY 958 GTGCGGCTTGAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATGTTGATGTCG 1017
DB 167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATACCTGTACAGTTTACTCCCTTAACCCGTATGACGATTTTCAGACGAGTCTCAAA 1077
DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1078 GATACCTAAGCTATTGAAAAACAATAAGTATGCGTACCCATGCATCTCAAGATTAATA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 1138 GCTCAAGCAACAAGCAATTTTAAACAACCAACCCGCGCTATACGATTTTGAACGCGAC 1197
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlyLysArg 246
QY 1198 TCCCTCAATCGTCACTGATGCAATGACATTTTCCGTACGATTTTCAACATGATCAAG 1257
DB 247 SerSerIleValIleHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
QY 1258 TTTTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCGTGTG 1317
DB 267 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnProLysThrGlyLe 286
QY 1318 AATGAAGAATAAACAACAACCTGACCTGATCTGAGAAATTTACGCTTAAAAAGCG 1377
DB 287 LysGlnLysThrAsnAsnThrAspLeuValSerGlnLysTyrTyrValLeuLysGlnGly 306
QY 1378 GAAAAAGCCGTATGATCCCTTGTGATCGAGTCACTTGAACTGTTCACATCAATAGCT 1437
DB 307 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGTCGATACCAACGAATGTCTAAAAAGTAGACAGCTCTTAAACAGCTACGGAACGTAC 1497
DB 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysArgAsn 346
QY 1498 TTAGACTTCAGAGATTATACGATCTCGTGAATAGGCTTAACCTTACCAACAATCTC 1557
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1558 GATGCTTTGGTATTATGACTATACCTTAACCTGAAAGATAGAGATATACGATGAC 1617
DB 367 AspAlaAspAspIleMetAspTyrThrLeuThrGlyLysValGlnLysAsnHisAspLys 386
QY 1618 ACCAACCGTATCATTAACCGTTTATATAGGCAAGCGAACCCGCAAGAGAAATGCTAGTAT 1677
DB 387 AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1678 CATTTAGCTTATGATTAAGAATCGTTATACCGAAGAAAGACGAGAACTTTACACTAC 1737
DB 407 HisLeuAlaTyrAspLysAspLeuTyrThrGlnGlnGlnLysAlaTyrSerTyrLeu 426
QY 1738 CATTATACAGGAGACACTTACTGATTAACCCCTAACGACAAA 1779

```

```

DB 427 ArgAspThrGlnGlyThrProIleProAspAsnProLysAspLys 440
RESULT 7
ID Q53ZK6_STRPY PRELIMINARY; PRT; 440 AA.
AC Q53ZK6;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Streptokinase.
GN Name=eka;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86-779;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalish A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations."
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234128; AAP39948.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;
Alignment Scores:
Pred. No.: 1,94e-126 Length: 440
Score: 1937.00 Matches: 367
Percent Similarity: 93.48% Conservative: 20
Best Local Similarity: 88.65% Mismatches: 27
Query Match: 61.81% Indels: 0
Gaps: 0
US-09-940-235-11 (1-1782) x Q53ZK6_STRPY (1-440)
QY 538 ATTGCTGACCTGATAGTGGCTGCTAGACCGGTCCATCTGTCAACAACGCAATTGGTGT 597
DB 27 IleAlaGlyTyrGlyTyrPheProAspArgProProIleAsnAsnSerGlnLeuVal 46
QY 598 AGCTTCCTGCTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 657
DB 47 SerMetAlaGlyIleValGlnGlyThrAspLysValPheIleAsnPheGlnIle 66
QY 658 GATTAACATCAACGACCTGCTCATGAGAAAGACAGACGAGCTTAAGTCCAAATCA 717
DB 67 AspleuThrSerGlnProAlaHisGlyGlySerThrGlnGlnGlyLeuSerProLysSer 86
QY 718 AAACCAATTGCTAGTGTAGTGGGAGATGCATTAACCTTGAAAGCGAATTCTTA 777
DB 87 LysProPheMetThrAspAsnGlyAlaMetProHisLysLeuGlyValAspLeuLeu 106
QY 778 AAGCTATTTCAGAAACAATTGATCGCTAACGTCACAGTAACGACGACTTACTTGAGGTC 837
DB 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 838 ATTGATTTTGCAAGCGATGCAACCATTAAGTGAAGCGGCAAGGTCTTACTTGCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 146
QY 898 AAAGATGTTGGTGAACCTTGCCGACCACTGTCACAAAGATTTTGGTGAAGGACAT 957
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuLysGlyHis 166
QY 958 GTGCGGCTTGAGACCATATAAAGAAAAACAATCAAAACCAAGGAAATGTTGATGTCG 1017
DB 167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATACCTGTACAGTTTACTCCCTTAACCCGTATGACGATTTTGAACGAGTCTCAAA 1077
DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206

```



```
QY 1078 GATACAACTATTGAAACACTAGCTATCGTGACACATCATCTCAAGATTACTA 1137
DB 207 AepThrvLeuLeuLeuYstrHleuAaIleGlyAspThrIleThrsErGlnGluLeu 226
QY 1138 GCTCAAGCAACAAGCTTTTAAACAACCAACCGAGCTATACGATTATGACGCTGAC 1197
DB 227 AAGlnAaIleGlnSerIleLeuLeuYstrHleuProGlyTyrThrIleTyrGluAAsp 246
QY 1198 TCCTCAATGTCATCATGACATGACATTTCCGACATTTTACCAATGATGACAGAG 1257
DB 247 SerSerIleValThrHlsAspAsnIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1258 TTACTTACCGTGTAAAATATCGGGAACAGCTTATAGATGATATAAATAATGGCTG 1317
DB 267 PheThrTyrHlsValIleValAspAsnGlnGlnAaIleTyrGluLeuAsnProLysTrnGlyIle 286
QY 1318 AATGAGAATAAACAACAACGACCTGATCTGTGAGAAATATTACGTCCTTAAAAAGG 1377
DB 287 LysGlnLysThrAspAsnThrAspLeuValSerGlnLysTyrTyrValLeuLysGlnGly 306
QY 1378 GAAAAGCGGTATGATCCCTTTGATGCGACGCTGACCTGAAACTGTTCAACATACGTT 1437
DB 307 GlnLysProLysAspProPheAspSerHlsLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGCGATACGACGAATTTGCTAAAAGTGAAGCAGCTCTTAAACAGCTAGCGGAACGTAAC 1497
DB 327 AspValAsnThrAspAsnIleLeuLeuYstrGlnGlnLeuThrLlsErGlnValArgAsn 346
QY 1498 TTAGACTTCAGAGATTTATACGATCCGTCGATGATAGGCTAAACTACTTACCAACATCTC 1557
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspValIleLysLeuLeuTyrAspAsnLeu 366
QY 1558 GATGCTTTTGGTATTATGACATATACCTTAACTGAAAAGTGAAGATATACGATGAC 1617
DB 367 AspAlaPheAspIleMetAspTyrThrLeuTrnGlyValGlnLysAsnHlsAspLys 386
QY 1618 ACCAACCAGTATACATACCGTTATATAGGCAAGCGACCGCAAGAGAAGATGCTACTAT 1677
DB 387 AsnAsnAlaGlnValThrValTyrMetGlyLysArgProLysGlnAlaLysGlnSerTyr 406
QY 1678 CATTAGCCTATGATTAAGATCGTTATACCGAAGAAGACGAGATTATACAGCTACCTG 1737
DB 407 HlsLeuAlaIleTyrAspLysAspLeuTyrTrnGlnGlnGlnLysAlaIleTyrSerTyrLeu 426
QY 1738 CGTTATACGAGGACACCTTATACCTGATAACCTTAAACGACAAA 1779
DB 427 ArgAspTrnGlyThrProIleProAspAsnProLysAspLys 440

RESULT 8
Q8K5R8_STRP3 PRELIMINARY; PRT; 440 AA.
ID Q8K5R8_STRP3 PRELIMINARY; PRT; 440 AA.
AC Q8K5R8; Q79W73;
RC STRAIN-MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
Mammarella N.D., Liu M.-Y., Smoot J.C., Forcella S.F., Parkins L.D.,
Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055 (2003).
DR EMBL; AB014169; AAM80305.1; -; Genomic DNA.
DR EMBL; BA000034; BAC64795.1; -; Genomic DNA.
DR HSSP; P00779; 114D.
DR SMK; Q8K5R8; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; Kinase.
SQ SEQUENCE 440 AA; 49911 MW; 7CCE844F4026E3975 CRC64;

Alignment Scores:
Pred. No.: 1,94e-126 Length: 440
Score: 1937.00 Matches: 372
Percent Similarity: 93.72% Conservative: 16
Best Local Similarity: 89.86% Mismatches: 26
Query Match: 61.81% Indels: 0
DB: Gaps: 0

US-09-940-235-11 (1-1782) x Q8K5R8_STRP3 (1-440)
QY 538 ATTGCTGACGCTGAGGCGCTGACGCGCTGATGTCACCAAGCAATGGTGT 597
DB 27 IleAlaGlyTyrGlyTyrLeuProAspArgProProValAsnAsnSerGlnLeuVal 46
QY 598 AGCGTGTGCTGACTGTGAGGAGGAGCAACAAGACTTATGCTTAAATTTTGAATC 657
DB 47 SerMetAlaGlyIleValGlnGlnGlnThrAspLysValPheIleAsnPheGlnIle 66
QY 658 GATCTAACATCAGCAGCTGCTCATGAGAGAAAGACAGAGAGCTTAACTCAAAATCA 717
DB 67 AspLeuThrSerGlnIleAlaHlsGlnGlyLysTyrThrGlnGlnGlnLysSerProLysSer 86
QY 718 AAACCATTTGCTACGTATGAGCGGCGGAGTCTCATTAATTTGAGAAAGCTGACTTA 777
DB 87 LysProPheAlaThrAspAsnGlnLysAlaMetProHlsLysLeuGlnLysAlaAspLeu 106
QY 778 AAGCTATTCAAGACAATGATGCTTACGCTTACGTCACAGTACAGCACTACTTTAGAGTC 837
DB 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHlsSerAsnAspGlyTyrPheGlnVal 126
QY 838 ATTGATTTTGCAGAGGATGACACCATTAATGATGAAACGGCAAGCTTACTTGTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlnLysValTyrPheAlaAsp 146
QY 898 AAAGATGCTTGGTATACCTTGGCCGACCCCAACCTGTCCAAGAAATTTTGTACCGGACAT 957
DB 147 LysAspGlySerValThrLeuProTrnGlnProValGlnGlnLysLeuLeuSerGlyHls 166
QY 958 GTGCGGTTTGAACATATAAAGAAACCAATCAAAACCAAGCAAGAAATGTTGATGTG 1017
DB 167 ValArgValArgProLysTyrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1018 GAATATACGTATACAGTTTACTCCCTTAAACCTGATGACGATTTACAGACAGGCTTCAAA 1077
DB 187 LysTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLys 206
QY 1078 GATCTAACGCTATTGAAAACAACGATGCTATCGTGACACCATCATCTCAAGATTACTA 1137
```



```

Db      |||
207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrsErgIngluLeu 226
QY      1138 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCCGAGGCTATGCAATTAATGAACGTGAC 1197
Db      227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisSerProGlyTyrThrIleTyrGluAsp 246
QY      1198 TCCTCAATCGTCACTCATGACATGACATTTTCCTGATGCAATTTTACCAATGATCAAGAG 1257
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY      1258 TTTACTTACCGGTGTTAAATATGGGAAACAAGCTTATAGATCAATTAATAAATCTGGCTG 1317
Db      267 PheThrTyrArgValLysAspArgGluGlnAlaTyrGlyIleAsnLysSerGlyLeu 286
QY      1318 AATGAAGAAATTAACAACATGACCTGATCTGTCGAGAAATATTAACGTCCTAAAGAGG 1377
Db      287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrIleLeuLysGly 306
QY      1378 GAAAAGCCGTATGATCCCTTGTATGACAGTCACTTGAACCTGTTCAACATCAATACGTT 1437
Db      307 GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
QY      1438 GATGTCGATACCAACGAATGCTTAAAGAGAGCAGCTCTTACACCTAGCGAAGCTAAC 1497
Db      327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY      1498 TTAGACTTCAGAGATTTATACATCTCTGATGATAGGCTTAACTCTTACCAATCAATCTC 1557
Db      347 LeuAspPheArgAspLeuTyrAspProCysAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY      1558 GATGCTTTGGTATATGATGACATATACCTTAACTGGAAGAGTGAAGATATACATGAC 1617
Db      367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
QY      1618 ACCAACCGTATCATACCGTTTATATGGGCAAGCGAACCGAAGAGAGATGCTACTAT 1677
Db      387 AsnAsnArgIleValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY      1678 CATTTAGCTTATGATTAAGATGTTATATCCGAAAGAAAGCAAGATTACAGCTACCTG 1737
Db      407 HisLeuAlaTyrAspLysAspLeuTyrThrGluGluGluArgLysAlaTyrSerTyrLeu 426
QY      1738 CGTTATACAGGAGACCTATACCTGATACCTTAACCTTAACGACAA 1779
Db      427 ArgAspThrGluThrProIleProAspAsnProLysAspLys 440

RESULT 9
Q7X0Y2_STRPY PRELIMINARY; PRT: 440 AA.
AC Q7X0Y2: 2003 (TREMBLrel. 25, Created)
RC PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.",
RL J. Bacteriol. 186:110-121(2004).
DR EMBL: AY234136; AAF39956.1; -; Genomic_DNA.
DR HSSP: Q53284; 1C4P.
DR SMR: Q7X0Y2; 63-398.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0008243; F:plasmidogen activator activity; IEA.

```

```

DR GO: GO:0005515; F:protein binding; IEA.
DR InterPro: IPR004093; Staphylokinase.
DR InterPro: IPR008124; Streptokinase.
DR Pfam: PF02821; Staphylokinase; 3.
DR PRINTS: PR01753; STREPKINASE.
DR Kinase.
SQ SEQUENCE 440 AA; 5018 MW; 4CB42AC2A9062C2E CRC64;

Alignment Scores:
Pred. No.: 5,96e-126 Length: 440
Score: 1930.00 Matches: 368
Percent Similarity: 93.72% Conservative: 20
Best Local Similarity: 88.89% Mismatches: 26
Query Match: 61.58% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y2_STRPY (1-440)
QY      538 ATTGCTGAACCTGAGTGGCTGCTAGACCGTCCATCTGTCACAAACAGCAATGGTGT 597
Db      27 IleAlaGlyTyrGlyTyrPleuProAspArgProIleAsnAsnSerGlnValVal 46
QY      598 AGCGTTCGTGATCTGTGAGGGAGCAATCAAGACATTAAGCTTAAATTTTGAATC 657
Db      47 SerMetAlaGlyIleValGluGlyThrAspLysValPheIleAsnPheGluIle 66
QY      658 GATCTTACATACGACCTGCTCATGAGGAAAGACAGGCAAGGCTTAAGTCCAAATCA 717
Db      67 AspLeuThrSerGlnProAlaHisGlyLysThrGluGlnGluLysSerProLysSer 86
QY      718 AAACATTTCTACTGATGAGGCGGCGATGTCACATTAACCTTGAAGAGCTTACTA 777
Db      87 LysProPheAlaThrAspAsnSerAlaMetProHisLysLeuGluLysAlaAspLeu 106
QY      778 AAGCTATTCAAGAAACATTAATGATCGTACGTCACAGTACGACTTCTTGAAGTGC 837
Db      107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY      838 ATTGATTTTCAAGCGATGCAACCATTAATCTGATGAAACGGCAAGTCTACTTGTGAC 897
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY      898 AAAGATGTTCCGTTAACCTTCCGACCCAGCAAGCTGCAAGATTTTGGAGGACAT 957
Db      147 LysAspLysSerValThrLeuProThrGlnProValGlnGluPheLeuThrGlyHis 166
QY      958 GTGCGGCTTGAACCATTAAGAAACCAATACAAACCAAGCAAGCAATCTGTGAGTGT 1017
Db      167 ValArgValArgProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY      1018 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGCAAGTCTGAA 1077
Db      187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY      1078 GATATTAAGCTATTTGAAACATCAGTATGCGTGACACCATCAATCTCAAGAAATTACTA 1137
Db      207 AspThrLysLeuLeuLysLysLeuAlaIleGlyAspThrValThrSerGlnGluLeu 226
QY      1138 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCCGAGGCTATGCAATTAATGAACGTGAC 1197
Db      227 AlaGlnAlaGlnSerIleLeuAsnGluSerHisProAspTyrThrIleTyrGluArgAsp 246
QY      1198 TCCTCAATCGTCACTCATGACATGACATTTTCCTGATGCAATTTTACCAATGATCAAGAG 1257
Db      247 SerSerIleValThrHisAspAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY      1258 TTTACTTACCGGTGTTAAATATGGGAAACAAGCTTATAGATCAATTAATAAATCTGGCTG 1317
Db      267 PheThrTyrArgValLysAsnAspArgGluGlnAlaTyrArgIleAsnLysSerGlyLeu 286
QY      1318 AATGAAGAAATTAACAACATGACCTGATCTGTCGAGAAATATTAACGTCCTAAAGAGG 1377
Db      287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrIleLeuLysGlyArg 306

```

QY 1378 GAAAACCGTATGATCCCTTGTATGCGAGTCACTTGAAACGTTCACATCAATAGTT 1437
 |||||
 Db 307 GIUlySProTyAspProTheAspArgSerHisLeuIySLeuPheThrIleTySVal 326
 QY 1438 GATGTGATACCAACGAATTTGCTAAAGTAGACAGCTTTAAACAGCTTACGAAAGCTAAC 1497
 |||||
 Db 327 AspValAsnThrAsnGlnLeuLeuIySerGlnGlnLeuLeuThrAlaSerGlnIyAsn 346
 QY 1498 TTAGACTTCAGAGATTATACGATCCCTGATTAAGCTTAACCTTCAACAACTTC 1557
 |||||
 Db 347 LeuAspPheArgAspLeuTyAspProArgAspIyAlaIySLeuLeuTyAsnAsnLeu 366
 QY 1558 GATGCTTTGTATTTGACTTATGACTTATACCTTATACGAAAGTAGAGATTAACGATGAC 1617
 |||||
 Db 367 AspAlaPheAspIleMetAspTyThrLeuThrGlyIySValGlnAspAsnHisAspIyS 386
 QY 1618 ACCAACCGTATCATTAACCGTTATATATGGCAACGACCCGAGAGAGAAATGCTAGCTAT 1677
 |||||
 Db 387 AsnAsnArgValAlaThrValIyMetGlyIyArgProIySgIyAlaIySgIySerTy 406
 QY 1678 CATTTAGCTTATGATTAAGATCGTTATACCGAAGAAAGAGAAAGTTTACAGTACCTG 1737
 |||||
 Db 407 HisLeuAlaTyAspIySAspLeuTyThrGlnGlnIyAspIySAlaTySerTyLeu 426
 QY 1738 CGTTATACAGGAGACACTTATACCTGATTAACCTTAAGCAAA 1779
 |||||
 Db 427 ArgAspThrGlyThrProIleProAspAsnProIySAspIyS 440

RESULT 10

Q7X0Y8_STRPY PRELIMINARY; PRT: 440 AA.

ID 07X0Y8_STRPY
 AC 07X0Y8-
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=88-019;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Bessen D.E.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL: AY234129; AAP3949.1; -; Genomic_DNA.
 DR HSP; Q53284; 1C4P.
 DR SMR; Q7X0Y8; 63-398.
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0008243; P:plasmaogen activator activity; IEA.
 DR GO; GO:0005515; P:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STRPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 49865 MW; 69DB4F4026B3975 CRC64;

Alignment Scores:

Pred. No.: 7e-126 Length: 440
 Score: 1929.00 Matches: 371
 Percent Similarity: 93.48% Conservative: 16
 Best Local Similarity: 89.61% Mismatches: 27
 Query Match: 61.53% Indels: 0
 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y8_STRPY (1-440)

QY 538 ATTGCTGACCTGATGCTGCTAGACCGTCCATCTGTCAACAAGCGCAATGTTGTT 597
 |||||
 Db 27 IleHisGlyTyArgIyTyrLeuProAspArgProIyValAsnAsnSerGlnLeuVal 46
 QY 598 AGCGTTGCTGACTGTGGGGGAGCAATCAAGACTTACTTTAAATTTTGAATC 657
 |||||
 Db 47 SerMetAlaGlyIleValGlnGlyThrAspIySValPheIleAsnPhePheGlnIle 66
 QY 658 GATTTACATACGACCGCTGATGCGGAAAGCAAGCAAGGCTTAAGTCAAAATCA 717
 |||||
 Db 67 AspLeuThrSerGlnHisAlaHisGlyIyLeuThrGlnGlnIyLeuSerProIySer 86
 QY 718 AAACATTTCTACTGATAGTGGCGCATGATCATTAACCTTGAAGAAAGCTTACTA 777
 |||||
 Db 87 TyrProHealThrAspAsnGlyAlaMetProHisIySLeuGlnIySAlaAspLeuLeu 106
 QY 778 AAGGCTATTCAGAAACATTAATGATCGTCAACGTCACAGTAACGACTTTGAGGTC 837
 |||||
 Db 107 LysAlaIleGlnIySgIleuIleAlaAsnValHisSerAsnAspGlyTyRheGlnVal 126
 QY 838 ATTGATTTTGCAGCGGATGCAACCATTAAGTGAAGAAAGGCTTACTTGTGAC 897
 |||||
 Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlySValTyRheAlaAsp 146
 QY 898 AAAGATGCTTCGGTAACTTGCCGACCAACCTGTCAAGAAATTTGCTTAAGCGGACAT 957
 |||||
 Db 147 LysAspIySerValThrLeuProThrGlnProValGlnGlnIyPheLeuSerGlyHis 166
 QY 958 GTGGCGTTAGACCATTAAGAAACCAATACAAACCAAGCAAGCAATCTTGTATGTC 1017
 |||||
 Db 167 ValArgValArgProTyRlySgIySProValGlnAsnGlnAlaIySValAspVal 186
 QY 1018 GAATATATCTGATAGTTTACTCCCTTAACCTGATGACGATTTCAAGACCGGTCTCAA 1077
 |||||
 Db 187 LysTyThrValGlnPheThrProLeuAsnProAspAspPheAspProGlyLeuIyS 206
 QY 1078 GATATCAAGCTATTAAGAAACATGATGCGTGAACCATCATCATGCTCAAGAAATTA 1137
 |||||
 Db 207 AspThrIySLeuLeuIySThrLeuAlaIleGlyAspThrIleThrSerGlnIyLeuLeu 226
 QY 1138 GCTCAAGCAACAAAGCATTTTAAACAAACCAACCCAGCTATACGATTTTGAACGTGAC 1197
 |||||
 Db 227 AlaGlnAlaGlnSerIleLeuAsnIySThrHisProGlyTyThrIleTyGlnIyAsp 246
 QY 1198 TCCTCAATGCTGACTCATGACATGACATTTTCCGATTCGATTTTCAAGTGAACGAC 1257
 |||||
 Db 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIy 266
 QY 1258 TTTACTACCGGTGTTAAATTCGGGAACAGCTTATGATCATTAATAATCGGTCTG 1317
 |||||
 Db 267 PheThrTyArgValIyAspAspArgGlnGlnAlaTyGlyIleAsnIySValSerGlyLeu 286
 QY 1318 AATGAAGAAATTAACACACTGACCTGATCTGTGAGAAATATATCGCTTTAAAGAGG 1377
 |||||
 Db 287 AsnGlnIyIleAsnAsnThrAspIleIleSerGlnIySValTyThrIleLeuIySVal 306
 QY 1378 GAAAACCGTATGATCCCTTGTATGCGAGTCACTTGAACSTGTTCACCATCAATAGCTT 1437
 |||||
 Db 307 GluSerProTyAspProPheAspArgSerHisLeuIySLeuPheThrIleTySVal 326
 QY 1438 GATGTGATACCAACGAATTTGCTAAAGTAGAGAGCTTTAAACAGCTTACGAAAGCTAAC 1497
 |||||
 Db 327 AspValAsnThrAsnGlnLeuLeuIySerGlnGlnLeuLeuThrAlaSerGlnIyAsn 346
 QY 1498 TTAGACTTCAGAGATTATACGATCCCTGATTAAGCTTAACCTTCAACAACTTC 1557
 |||||
 Db 347 LeuAspPheArgAspLeuTyAspProArgAspIyAlaIySLeuLeuTyAsnAsnLeu 366
 QY 1558 GATGCTTTGTATTTGACTTATGACTTATACCTTATACGAAAGTAGAGATTAACGATGAC 1617
 |||||
 Db 367 AspAlaPheAspIleMetAspTyThrLeuThrGlyIySValGlnAspAsnHisAspIyS 386
 QY 1618 ACCAACCGTATCATTAACCGTTATATATGGCAACGACCCGAGAGAGAAATGCTAGCTAT 1677

```

Db      387 AanaAsnArgIleValThrValIlyrMeGlyLysArgProLysGlyAlaLysGlySerTyr 406
Qy      1678 CATTAGCCTATGATTAAGAATCGTTATACCGAAGAAAGAAAGAGTTTACAGTACTG 1737
Db      407 HisLeuAlaIlyrAspLysAspLeuIlyrThrGluGluGluIlyrLysAlaIlyrSerTyrLeu 426
Qy      1738 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAA 1779
Db      427 ArgAspThrGluThrAlaIleProAspAsnProLysAspLys 440

RESULT 11
Q7X0Y7_STRPY PRELIMINARY; PRT; 440 AA.
ID Q7X0Y7_STRPY
AC Q7X0Y7;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=Bka;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus;
OC NCB1_TaxID=1314;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IRP112;
RX PubMed=14679231; DOI=10.1126/JB.186.1.110-121.2004;
RA Kalia A., Beesen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
   involved in tissue-specific adaptations.";
RT J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234130; AAP3950.1; -; Genomic_DNA.
DR HSSP; 053284; 1CAP.
DR SMR; Q7X0Y7; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PRO1753; STREPKINASE.
DR KINASE.
KV SEQUENCE 440 AA; 50160 MW; 1CSAF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 1,13e-125 Length: 440
Score: 1926.00 Matches: 368
Percent Similarity: 93.48% Conservative: 19
Best Local Similarity: 88.89% Mismatches: 27
Query Match: 61.46% Indels: 0
                Gaps: 0
                2

US-09-940-235-11 (1-1782) x Q7X0Y7_STRPY (1-440)
Qy      538 ATTGCTGAGCCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACGCAATGTTGTTT 597
Db      27 IleaIaGlyTyrGlyTyrLeuLeuAspArgProProValaAsnAsnSerGlnLeuVala 46
Qy      598 AGCGTTGCTGTAAGTCTTGAAGGGAAGCAATCAAGATTAAGTCTTAAATTTTGAATC 657
Db      47 SerMetAlaGlyIleValaGluGlyThrAspLysLysValaPheIleAsnPhePheGluIle 66
Qy      658 GATTAACATCAAGACCTGCTCAAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
Db      67 AspLeuThrSerGlnProAlaHisGlyLysThrGluGlnIlyLeuSerProLysSer 86
Qy      718 AAACCATTTGCTACTGATAGTGGCGGAGTGTCACATTAATTGAGAAAGTGACTTA 777
Db      87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeu 106
Qy      778 AAGGCTATTCAGAGAACATTAATGATCGTCAAGTCCAGAGTACGACTTGAAGTTC 837

```

```

Db      107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVala 126
Qy      838 ATTGATTTTGGACAGATGCAACCATTAATGATCGAAACGGCAAGGTGTACTTTGCTAC 897
Db      127 ILeAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValaTyrPheAlaAsp 146
Qy      898 AAAGATGATTCGGTACCTTGCCGACCCCAACCTGCTCCAAAGAAATTTTGTCAAGCGACAT 957
Db      147 LysAspSerSerValThrLeuProThrGlnProValGlnGluPheLeuLeuAlaGlyHis 166
Qy      958 GTGCGCGTTAGACCATTAATAAGAAAACCAATACAAAACCAAGCAAAATCTGTGATGTG 1017
Db      167 ValArgValArgProTyrLysGluLysProIleGlnThrProAlaLysSerValaAspVal 186
Qy      1018 GAATATACGTGACGTTTACCTCCCTTAACCGCTGATGACGATTTTCAAGCGCTCAAA 1077
Db      187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheAspProValLeuLys 206
Qy      1078 GATACATAAGCTATTGAAAACATGACTACGCTGACACATCAATCTCAAGAAATTA 1137
Db      207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyLysThrValThrSerGlnGluLeu 226
Qy      1138 GCTCAGACACAAAGCATTTTAAACAAAACCAAGCGCTATACGATTATGAACGTGAC 1197
Db      227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisAspArgTyrThrIleTyrGluArgAsp 246
Qy      1198 TTCCTCATCGTCACTGACATGACAAATGACATTTCCGTACGATTTTACCAATGATCAAGAG 1257
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy      1258 TTTACTACCGTGTAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCGTGCTG 1317
Db      267 PheThrTyrThrIleLysAspArgGlnGlnAlaTyrGlyIleAsnLysSerGlyGln 286
Qy      1318 AATGAGAAATTAACCAACCTGACCTGATCTTGAGAAATTAAGTCTTAAAAAGGG 1377
Db      287 GlnGluLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
Qy      1378 GAAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTCACCATCAATACGTT 1437
Db      307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
Qy      1438 GATGCTGATCAACCAATGCTTAAAAAGTACAGGCTTTAACAGCTACGACGAAACGTAAC 1497
Db      327 AspValAsnThrIleAsnLysLeuLysSerGlnIleLeuLeuThrAlaSerGlnArgAsn 346
Qy      1498 TTAGACTTCAGAGATTTATACGATCTCGTATAGGCTTAACTTACTCAACAACTCTC 1557
Db      347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
Qy      1558 GATGCTTTGGTATTAAGACATACTTAACTGGAAGAGTGAAGATTAATCACGATGAC 1617
Db      367 AspAlaPheGlyIleMetAspArgTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
Qy      1618 ACCAACCGTATCAATTAACGTTTATATAGGGAAGCAAGCAAGCAAGCAAGTATGAT 1677
Db      387 AsnAsnArgIleValThrValIlyrMeGlyLysArgProGluGluIlyAsnAlaSerTyr 406
Qy      1678 CATTAGCCTATGATTAAGAATCGTTATACCGAAGAAAGAAAGAGTTTACAGTACTG 1737
Db      407 HisLeuAlaIlyrAspLysAspArgTyrThrGluGluGluIlyrLysAlaIlyrSerTyrLeu 426
Qy      1738 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAA 1779
Db      427 ArgTyrThrGlyThrProIleProAspAsnProLysAspLys 440

RESULT 12
Q5X9T6_STRP6 PRELIMINARY; PRT; 440 AA.
ID Q5X9T6_STRP6
AC Q5X9T6;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)

```

DE Streptokinase (EC 3.4.-.-).
 GN OrderedLocusNames=M6_Spy1692;
 OS Streptococcus pyogenes (serotype M6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCB1_TaxID=301450;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MGAS10394;
 RX PubMed=15272401; DOI=10.1086/422697;
 RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.R.,
 RA Voyich J.M., Deleo F.R., Martin J.M., Somerville G.A., Musser J.M.;
 RT "Progress toward characterization of the group A Streptococcus
 RT metagenome: complete genome sequence of a macrolide-resistant serotype
 RT M6 strain.";
 RL J. Infect. Dis. 190:727-738(2004).
 DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
 DR SMR; OSX9T6; 63-398.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasma/nogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KM Complete proteome; Hydrolase; Kinase.
 SQ SEQUENCE 440 AA; 50160 MW; 1CSAF07907ECTAC8 CRC64;

Alignment Scores:
 Pred. No.: 1,13e-125 Length: 440
 Score: 1926.00 Matches: 368
 Percent Similarity: 93.48% Conservative: 19
 Best Local Similarity: 88.89% Mismatches: 27
 Query Match: 61.46% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x OSX9T6_STRP6 (1-440)

QY 538 ATTGTGACCTGAGTGGCTGTGCTAGACCGTCATCTGTCAACACGCAATGTTGTT 597
 DB 27 IleaIaGIyTYrGlyTTrpleuLeuAaPaRgPrOProValaAsnaBserGInleuVal 46
 QY 598 AGCGTGTGCTGACCTGTGAGGGAGACGATCAAGACATTAGCTTAAATTTTGAATC 657
 DB 47 SerMetAlaGIyIleValGIuGIyThraSpLySLyValPheIleasnhePheGIuIle 66
 QY 658 GATCTAACATCAACGACCTGCTCATGAGGAGAAAGACAGACAGGCTTAAGTCAAAATGA 717
 DB 67 AspLeuIntrSerGIaProAlaHISGIyGIySLyThnGIuGIuGIyLeuSerProLySer 86
 QY 718 AAACCATTTGCTACTGATAGTGGCGCGATGTCACTAACTTGAGAACTGACTTACTA 777
 DB 87 LysPProPheAlaThraAsnSerSerAlaMetProHISLyLeuGIuLySLyAlaAspLeu 106
 QY 778 AAGGCTATTCAAGAACTATTGATCGCTAACGTCACAGTACGACGACTTGTGAGGTC 837
 DB 107 LysAlaIleGIuGIuGIyThraSpLySLyValPheIleasnhePheGIuIle 126
 QY 838 ATTGATTTTGAAGGAGATCAACCTTACTGATCGAAAGGAGGAGGCTTACTTGTGTCAC 897
 DB 127 IleAspPheAlaSerAspAlaThraIleIntraSpArgAsnGIySLyValIlyrPheAlaAsp 146
 QY 898 AAAGATGGTTCGGTAACTTGGCGACCCAACTGTCTCAAGAAATTTTGTAAAGCGACAT 957
 DB 147 LysAspAspSerValThraLeuProThraInProValaGIuGIuPheLeuLeuArgGIyHIS 166
 QY 958 GTGGCGCTTGAACATATTAAGAAAAACAATACAAACAAAGCAAGCAATCTGTTGATGTG 1017
 DB 167 ValArgValaArgProLyTrLySGIySLyProIleGIuThraProAlaLySerValaAspVal 186
 QY 1018 GAATTAAGTGAAGATTTACTCCCTTAAACCTGATGAGATTTTGACAGAGCTTCGAA 1077

DB 187 ArgTYrThraValGIuInPheThraProLeuAsnProAspAspPheArgProValaLeuLys 206
 QY 1078 GATACCTAAGCTATTGAAAAACATAGCTATGGTGACACCACTACATCTTCAAGAAATTTACTA 1137
 DB 207 AsnThraLyLeuLeuLysThraIleAlaIleGIyGIyThraValThraSerGIuGIuLeuLeu 226
 QY 1138 GCTCAAGCAACAAAGCAATTTTAAACAAAACACACCGAGGCTATACGATTTTGAACGTGAC 1197
 DB 227 AlaGIuInAlaInserIleLeuAsnGIuSerHISProAspTYrThraIleTYrGIuArgAsp 246
 QY 1198 TCCTCAATCGTCACTGATGACAAATGACATTTTCGTCAGATTTTACCAATGATGACAG 1257
 DB 247 SerSerIleValThraHISAspAsnAspIlePheArgThraIleLeuProMetAlaSpGIuGIu 266
 QY 1258 TTTACTTACCGTGTTTAAAAATCGGGAACAGCTTATAGATCAATATAAAATCTGCTCTG 1317
 DB 267 PheThraThraIleLysAspArgGIuGIuAlaTYrIleAsnLySLySerGIyGIu 286
 QY 1318 AATGAAGAAATAAACAACCTGACCTGATCTGAGAAATTTACGTCCTTAAAAAGGG 1377
 DB 287 GIuGIuLySLyThraAsnThraAspLeuIleSerGIuLySLyTYrValaLeuLySLyGIy 306
 QY 1378 GAAACCCGATGATGATCCCTTGTGATCGCAGTCACTTGAACGTTCACATCAATACGTT 1437
 DB 307 GIuLyProLyThraProPheAspArgSerHISLeuLySLyPheThraIleAsnTYrVal 326
 QY 1438 GATGTGATACCAACGATTTGCTTAAAGTAGAGAGCTTTTAAACAGCTTACGACGTAAC 1497
 DB 327 AspValaAsnThraAsnLySLyLeuLySLySerGIuGIuLeuThraIleAsnGIuArgAsn 346
 QY 1498 TTGACCTCAGAGATTTATATACGATCCCGGATGATGAGGCTTAACTACCTTCAACAACTCTC 1557
 DB 347 LeuAspPheArgAspLeuTYrAspProArgAspLySLyAlaLySLyLeuTYrAsnAsnLeu 366
 QY 1558 GATGCTTTTGTATATATGACACTATACCTTAACTGAGAAATAGAGATTAATCAAGATAC 1617
 DB 367 AspAlaPheGIyIleMetAspTYrThraLeuThraGIySLyValGIuAspAsnHISAspLySLy 386
 QY 1618 ACCAACCGTATCATTAACCGTTTATATGAGGACAGACCCGAGAGAGAAATGCTAGCTAT 1677
 DB 387 AsnAsnArgValaValaThraValaTYrMetGIyLysArgProGIuGIuAsnAlaSerTYr 406
 QY 1678 CATTAGCTATGATTAAGATCGTTATACCGAAGAAAGAAAGATTTACAGTACTCTG 1737
 DB 407 HISLeuAlaIlyThraSpLyAspArgTYrThnGIuGIuGIuValaIlySerTYrLeu 426
 QY 1738 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAA 1779
 DB 427 ArgTYrThraGIyThraProIleProAspAsnProLyAspLySLy 440

RESULT 13
 Q8NZAG_STRP8
 ID Q8NZAG_STRP8 PRELIMINARY; PRT; 440 AA.
 AC Q8NZAG6.
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 GN Streptokinase.
 OS OrderedLocusNames=spyM18_2042;
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCB1_TaxID=301451;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MGAS8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 RT group A Streptococcus strains associated with acute rheumatic fever

RT outbreaks.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
 DR EMBL; A6010107; AAL98517.1; -; Genomic_DNA.
 DR HSSP; O53284; 1C4P.
 DR SMR; O8NZ6; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 DR Complete proteome; Kinase.
 SQ SEQUENCE 440 AA; 50382 MW; 47DAB3557EBB88B CRC64;

Alignment Scores:

Pred. No.:	1.83e-125	Length:	440
Score:	1923.00	Matches:	364
Percent Similarity:	93.96%	Conservative:	25
Best Local Similarity:	87.92%	Mismatches:	25
Query Match:	61.36%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-11 (1-1782) x Q8NZ6_STRP8 (1-440)

QY 538 ATGTGACGCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATGGTGT 597
 DB 27 TLeaIGlYrGtUtrPLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
 QY 598 AGCGTGTCTGTACTGTGAGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATC 657
 DB 47 SerMetAlaGlyLeuValGlnGlyThrAspLeuValPheIleAsnPheGlnIle 66
 QY 658 GATCTAACATCAGACCTGTCTATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 717
 DB 67 AspLeuThrSerIleProAlaHisGlyLysThrGlnGlnGlyLeuSerProLysSer 86
 QY 718 AAACGATTTGCTAGTCTAGTGGCGGATGTCACATTAACCTTGAAGGATCTTCTA 777
 DB 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlyValAspLeuLeu 106
 QY 778 AAGGCTATTCAAGACATTTGATGCTTACGTCACAGTACGACGACTTTGAGTCT 837
 DB 107 LysAlaIleGlnIleArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
 QY 838 ATTGATTTTGCAGACGATGCAACCATTAAGTCTGATGAAAGCGCAAGTCTTAC 897
 DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
 QY 898 AAAGATGTTCCGCTAACCTTGCAGACCCACCTGTCCAAAGATTTTGGTAAAGCAT 957
 DB 147 GlnAspGlySerValThrLeuProIleGlnProIleGlnIlePheLeuLeuAlaGlyHis 166
 QY 958 GTGCGCGTTAGACCATTAAGAAACCAATACAAACCAAGGGAATGTTGATGATG 1017
 DB 167 ValArgValArgProTyrLysGlyLysProIleGlnThrProAlaLysSerValAspIle 186
 QY 1018 GAATATACGTATACGTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGTCTCAA 1077
 DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValIleLys 206
 QY 1078 GATACCTAGCTATTGAAACCACTAGCTATGGTGACCACTACATCTCAAGATTTACTA 1137
 DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
 QY 1138 GGTCAAGCAACAAAGCATTTTAAACAAACCAACCAAGCGCTATCGATTTTGAACGTCAC 1197
 DB 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisSerAspTyrThrIleTyrGluArgAsp 246
 QY 1198 TCCCTCAATGCTCATCATGACATGACATTTTCGTCAGATTTTTCACAGATGATGAG 1257
 DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266

QY 1258 TTTACTTACCGGTGTTAAAAATCGGAGACAGCTTATAGGATCATATAAAAAATCTGCTG 1317
 DB 267 PheThrTyrHisIleLysAspArgGlnGlnIleTyrGlyIleAsnLysLysSerIleGln 286
 QY 1318 AATGAGAAATTAACACACATGACCTGATCTCTGGAATATATTACGCTCTTAAAAAGG 1377
 DB 287 GlnGlyLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
 QY 1378 GAAAGCCGCTATGATCCCTTTGATCGGACGCTACTGGAACCTGTTCACATCAAACTGTT 1437
 DB 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
 QY 1438 GATGTCATACCAAGCAATGCTTAAAGTAGACAGCTCTTACAGCTAGCAAGCTAAC 1497
 DB 327 AspValAsnThrAsnLysLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 346
 QY 1498 TTGACCTTCAGAGATTTATACGATCTGTGATTAAGGCTAACTACTACCAACATCTC 1557
 DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
 QY 1558 GATGCTTTGCTATTTATGACTTATACCTTAAGTGAAGAGTATGAGATTAATCAGATGAC 1617
 DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspLys 386
 QY 1618 ACCAACCGTATCATACCGCTTATATGAGGACGACCCGAGAGAGAGAAATGCTAGCTAT 1677
 DB 387 AsnAsnArgValIleThrValTyrMetGlyLysArgProGlnGlyLysAsnAlaSerTyr 406
 QY 1678 CATTAGCCTATGATTAAGATCGTTATACCGAAGAGACGAGAAAGATTGACGCTACCTG 1737
 DB 407 HisLeuAlaTyrAspLysAspArgTyrThrGlnGlnGlnValTyrSerTyrLeu 426
 QY 1738 CGTTATACAGGACGACTTATACCTGATACCTTAACCTTAACGACAA 1779
 DB 427 ArgTyrThrGlyThrProIleProAspAsnProLysAspLys 440

RESULT 14

Q7X0Y3_STRPY PRELIMINARY; PRT; 440 AA.
 ID Q7X0Y3;
 AC Q7X0Y3;

DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Streptokinase.

GN Name=aka;

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1314;

RX [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ALAB49;

RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;

RA Kalia A., Beasen D.B.;

RT "Natural selection and evolution of streptococcal virulence genes

involved in tissue-specific adaptations.";

RL EMBL; AY234134; AAP39954.1; -; Genomic_DNA.

DR HSSP; O53284; 1C4P.

DR SMR; O7X0Y3; 63-398.

DR GO; GO:0016301; F:kinase activity; IEA.

DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.

DR GO; GO:0005515; F:protein binding; IEA.

DR InterPro; IPR004093; Staphylokinase.

DR InterPro; IPR008124; Streptokinase.

DR Pfam; PF02821; Staphylokinase; 3.

DR PRINTS; PR01753; STREPKINASE.

SQ SEQUENCE 440 AA; 50342 MW; B3C6DBAA50DB2C39 CRC64;

Alignment Scores:

Pred. No.:	1.82e-122	Length:	440
Score:	1880.00 <td>Matches:</td> <td>356</td>	Matches:	356

Percent Similarity: 92.51% Conservative: 27
 Best Local Similarity: 85.99% Mismatches: 31
 Query Match: 59.99% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y3_STRPY (1-440)

```

QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGT 597
DB 27 ILeAlAGlyTyrGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 46
QY 558 AGCGTTGCTGTACTGTGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 657
DB 47 SerMetAlAGlyIleValGlnGlyThrAspIleValPheIleAsnPhepegIuIle 66
QY 658 GATCTTAACATCAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAACTCAAAATCA 717
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyIleThrGlnGlnGlyLeuSerProIleSer 86
QY 718 AAACCATTTGCTACTGATAGTGGCGAGTGTCAATTAACCTTGAGAAGCTGACTTA 777
DB 87 LysProPheAlaThrAspAsnGlyAlaMetProIleLysLeuGlnLysAlaAspLeuLeu 106
QY 778 AAGGCTATTCAAGAACAAATTGATGCTTAACTTCCAGAGTAACTGCTTGAAGTC 837
DB 107 LysAlaIleGlnGlnLysLeuIleAlaAsnValHisSerAsnAspGlyIleThrPheGlnVal 126
QY 838 ATTGATTTTGGCAAGCATGACCAACATTACTGATCAAGAACGCGACTTACTTGTCTGAC 897
DB 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleIleThrPheAlaAsn 146
QY 898 AAAGATGTTGGGTACTTGGCCGACCCCAACCTGTCCAGAAATTTTGTCAACGCGACAT 957
DB 147 GlnAspGlySerValThrLeuProIleGlnProIleGlnGlnPheLeuLeuAlaGlyHis 166
QY 958 GTGGCGGTGTGACCATATATAAGAAAAACAATCAAAACCAAGCAAAATCTGTTGATGTG 1017
DB 167 ValAlaGlyValArgProIleIleGlnLysProIleGlnIleThrProAlaLysSerValAspIle 186
QY 1018 GAATTAATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAGGCTTCAAA 1077
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspAspPheLysProValLeuLys 206
QY 1078 GATTAATCACTTGTGAACAACCTAGTATCGGTGACACATCAATCTCAAGAAATTACTA 1137
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 1138 GCTCAAGACCAAAAGCATTTTAAACAACAAACCCAGGCTATGCAATTTTGAACGCTGAC 1197
DB 227 AlaGlnIleGlnIleSerIleLeuAsnGlnIleSerHisProAspTyrThrIleTyrGlnLysAsp 246
QY 1198 TCCTCAATCGTCACTGATGACCAATGACATTTTCGATGATTTTACCAATGATCAAGAG 1257
DB 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProThrAspGlnGlu 266
QY 1258 TTTACTTACCGTGTAAATAATGGGGAACAAGCTTATAGATCAATAATAATCTGCTG 1317
DB 267 PheThrTyrHisIleValLysAsnArgGlnGlnAlaIleGlnLysAsnAspAsnIleThrGlnLysLeu 286
QY 1318 AATGAAGAATAAACAACACTGACCTGATCTGTGAATAATATATACCTTAAATAAGGG 1377
DB 287 LysLysGlnIleThrLysAsnThrAspLeuIleSerGlnLysTyrIleLeuLysLysGly 306
QY 1378 GAAAAGCCGTATGATCCCTTGTATCGACATCTTGAACCTGTCAACATCAATACGTT 1437
DB 307 GlnLysProIleAspProPheAspArgSerHisIleLeuLysLeuPheThrIleLysTyrVal 326
QY 1438 GATGTTCGATACCAAGCAATTTGCTAAATAAGTAGACAGCTTAAAGCTTACGACGACTAAC 1497
DB 327 AspValAspThrLysAspLeuLysSerGlnGlnIleLeuLeuThrAlaSerGlnLysAsn 346
QY 1498 TTAGACTTCAGAGATTATACGATCTCGTATTAAGCTTAACTACTCTTAACAATATTC 1557

```

```

DB 347 LeuAspPheArgAspLeuIleIleAspProArgAspIleAlaLysLeuLeuIleIleIle 366
QY 1558 GATGCTTTGGTATTATGACCTAATCACTTAACTGGAAGAGAGATTAATCAAGATGAC 1617
DB 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
QY 1618 ACCAACCGTATCATTAACCGTTTATATGAGGAGAGGACCCGACAGAGAGAAATGCTAGCTAT 1677
DB 387 AsnAsnArgValValThrValIleThrMetGlyLysArgProIleLysGlnLysGlySerTyr 406
QY 1678 CATTAGCTTATGATTAAGATCGTTATATCCGAAAGAACAGAGAGATTACAGCTACTG 1737
DB 407 HisLeuAlaIleIleAspIleAspLeuIleIleThrGlnGlnGlnLysGlnValIleIleSerTyrLeu 426
QY 1738 CGTTATACAGGACGACCTTATACCTGATTAACCTTAACGACGAA 1779
DB 427 ArgAspThrGlyThrProIleProAspAsnProAspAspLys 440

RESULT 15
Q7X0Y0_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=d633;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalia A., Bessen D.B.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DB EMBL; AY234138; AAP39958.1; -; Genomic_DNA.
DR HSSP; Q53284; 1CAP.
DR SMR; Q7X0Y0; 63-398.
DR GO; GO:0016301; F-kinase activity; IEA.
DR GO; GO:0008243; F-platelet activator activity; IEA.
DR GO; GO:0005515; F-protein binding; IEA.
DR InterPro; IPR004093; Streptokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PRO1753; STREPKINASE.
KW kinase.
SQ SEQUENCE 440 AA; 50143 MW; 4ACB2C29349D680C CRC64;

Alignment Scores:
Pred. No.: 6,58e-122 Length: 440
Score: 1872.00 Matches: 354
Percent Similarity: 92.03% Conservative: 27
Best Local Similarity: 85.51% Mismatches: 33
Query Match: 59.73% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-11 (1-1782) x Q7X0Y0_STRPY (1-440)

QY 538 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTGGTTGT 597
DB 27 ILeAlAGlyProGluTrpLeuLeuLysArgProValAsnAsnSerGlnLeuValVal 46
QY 598 AGCGTTGCTGTACTGTGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 657
DB 47 SerMetAlAGlyIleValGlnGlyThrAspIleValPheIleAsnPhepegIuIle 66
QY 658 GATCTTAACATCAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAACTCAAAATCA 717
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyIleThrGlnGlnGlyLeuSerProIleSer 86

```


QY 947 CAAAGATGCTTCGTAACCTTCCGACCCCACTGTCCAGATTTTCTTAAGCGACA 1006
DB 1256 CAAAGATGCTTCGTAACCTTCCGACCCCACTGTCCAGATTTTCTTAAGCGACA 1315
QY 1007 TGTGCGGTTAGACCAATATTAAGAAAACCAATACCAAGCGAAATCTGTGATGT 1066
DB 1316 TGTGCGGTTAGACCAATATTAAGAAAACCAATACCAAGCGAAATCTGTGATGT 1375
QY 1067 GGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGATTTAGACAGGCTCAA 1126
DB 1376 GGAATATCTGTACAGTTTACTCCCTTAAACCTGTATGAGATTTAGACAGGCTCAA 1435
QY 1127 AGATATTAAGCTATTGAAAACTAGCTATCGGTGACCAATCACTCTCAAGATTAAT 1186
DB 1436 AGATATTAAGCTATTGAAAACTAGCTATCGGTGACCAATCACTCTCAAGATTAAT 1495
QY 1187 AGCTCAAGCAAAAGATTTTAAACAAACCAACCGGCTATACGATTTATGAAAGCTGA 1246
DB 1496 AGCTCAAGCAAAAGATTTTAAACAAACCAACCGGCTATACGATTTATGAAAGCTGA 1555
QY 1247 CTCCTCAATCTGTACCTATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1306
DB 1556 CTCCTCAATCTGTACCTATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1615
QY 1307 GTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1366
DB 1616 GTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1675
QY 1367 GAATGAAGAAATTAACCAACTGACCTGATCTGTGAAATATTAAGTCTTTAAAAAGG 1426
DB 1676 GAATGAAGAAATTAACCAACTGACCTGATCTGTGAAATATTAAGTCTTTAAAAAGG 1735
QY 1427 GGAAGAGCGTATGATCCCTTGTATGAGTCACTTGAATCTGTGCAATCAATAGT 1486
DB 1736 GGAAGAGCGTATGATCCCTTGTATGAGTCACTTGAATCTGTGCAATCAATAGT 1795
QY 1487 TGAATGCAATCAACCAATGCTTAAAGTGAAGCTCTTAAACAGTACGCAAGCTAA 1546
DB 1796 TGAATGCAATCAACCAATGCTTAAAGTGAAGCTCTTAAACAGTACGCAAGCTAA 1855
QY 1547 CTTAGACTTCAAGATTTATAGATCTCTGTATTAAGGCTTAACTACTTCAACCAATCT 1606
DB 1856 CTTAGACTTCAAGATTTATAGATCTCTGTATTAAGGCTTAACTACTTCAACCAATCT 1915
QY 1607 CGATGCTTTGGATTAATGACCTATACCTTAAAGTGAAGATTAATCAAGTGA 1666
DB 1916 CGATGCTTTGGATTAATGACCTATACCTTAAAGTGAAGATTAATCAAGTGA 1975
QY 1667 CACCAACCGTATCATACCGTTTATATGCGCAAGCGAACCGAAGAGAGATGCTAGCTA 1726
DB 1976 CACCAACCGTATCATACCGTTTATATGCGCAAGCGAACCGAAGAGAGATGCTAGCTA 2035
QY 1727 CCATTTAGCTGTGCT 1742
DB 2036 TCATTTAGCTTATGAT 2051

RESULT 3
A04926 2568 bp DNA linear PAT 15-JUL-1993
LOCUS S.equisimilis shk gene for streptokinase.
DEFINITION A04926
ACCESSION A04926.1 GI:412219
VERSION A04926.1
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2568)
REFERENCE
AUTHORS Ferretti, J.J. and Malke, H.
TITLE Streptokinase-coding recombinant vectors
JOURNAL Patent: EP 015137-A 1 14-AUG-1985;
THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA

FEATURES
Source Location/Qualifiers
1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"

ORIGIN
Query Match 54.7%; Score 1146.4; DB 6; Length 2568;
Best Local Similarity 99.5%; Pred. No. 1.2e-297;
Matches 1150; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 587 TATTGCTGAGCTGAGTGTCTAGACCGTCACTGTGCAACAACAGCAATGTTGT 646
DB 896 TATTGCTGAGCTGAGTGTCTAGACCGTCACTGTGCAACAACAGCAATGTTGT 955
QY 647 TAGGCTTCTGTACTGTGAGGGGACCAATCAAGCACTTACTTTAAATTTTGAAT 706
DB 956 TAGGCTTCTGTACTGTGAGGGGACCAATCAAGCACTTACTTTAAATTTTGAAT 1015
QY 707 CGATCTAACATCAGGACCTGCTCATGAGGAAAGCAGAGCAAGGCTTAAGTCCAAATC 766
DB 1016 CGATCTAACATCAGGACCTGCTCATGAGGAAAGCAGAGCAAGGCTTAAGTCCAAATC 1075
QY 767 AAAACCAATTTGCTACTGATAGTGGCGCGATGTCACTTAACTTGAAGAGCTTACT 826
DB 1076 AAAACCAATTTGCTACTGATAGTGGCGCGATGTCACTTAACTTGAAGAGCTTACT 1135
QY 827 AAAGGCTATTCAGAAACAATGATGCTTACGCTCAAGTACAGTACGACGCTA 886
DB 1136 AAAGGCTATTCAGAAACAATGATGCTTACGCTCAAGTACAGTACGACGCTA 1195
QY 887 CATTGATTTTGGCAAGCCATGCAACCATTACTGATTCGAAACGGCAAGGCTTACTTGTCTGA 946
DB 1196 CATTGATTTTGGCAAGCCATGCAACCATTACTGATTCGAAACGGCAAGGCTTACTTGTCTGA 1255
QY 947 CAAAGATGCTTCCGTTAACTTCCGACCCCACTGTCCAAAGATTTTGTCTAAGCGGACA 1006
DB 1256 CAAAGATGCTTCCGTTAACTTCCGACCCCACTGTCCAAAGATTTTGTCTAAGCGGACA 1315
QY 1007 TGTGCGGCTTATGACCATATTAAGAAAACCAATCAAAAACCAAGGAAATCTGTGATGT 1066
DB 1316 TGTGCGGCTTATGACCATATTAAGAAAACCAATCAAAAACCAAGGAAATCTGTGATGT 1375
QY 1067 GGAATATCTGTACATTTTACTCCCTTAAACCTGTATGAGATTTTCAAGAGCTCTCAA 1126
DB 1376 GGAATATCTGTACATTTTACTCCCTTAAACCTGTATGAGATTTTCAAGAGCTCTCAA 1435
QY 1127 AGATATTAAGCTATTGAAAACTAGCTATCGGTGACCAATCACTCTCAAGATTAAT 1186
DB 1436 AGATATTAAGCTATTGAAAACTAGCTATCGGTGACCAATCACTCTCAAGATTAAT 1495
QY 1187 AGCTCAAGCAAAAGATTTTAAACAAACCAACCGGCTATACGATTTATGAAAGCTGA 1246
DB 1496 AGCTCAAGCAAAAGATTTTAAACAAACCAACCGGCTATACGATTTATGAAAGCTGA 1555
QY 1247 CTCCTCAATCTGTACCTATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1306
DB 1556 CTCCTCAATCTGTACCTATGACCAATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1615
QY 1307 GTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1366
DB 1616 GTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGTCT 1675
QY 1367 GGAATGAAGAAATTAACCAACTGACCTGATCTGTGAAATATTAAGTCTTTAAAAAGG 1426
DB 1676 GGAATGAAGAAATTAACCAACTGACCTGATCTGTGAAATATTAAGTCTTTAAAAAGG 1735
QY 1427 GGAAGAGCGTATGATCCCTTGTATGAGTCACTTGAATCTGTGCAATCAATAGT 1486
DB 1736 GGAAGAGCGTATGATCCCTTGTATGAGTCACTTGAATCTGTGCAATCAATAGT 1795
QY 1487 TGAATGCAATCAACCAATGCTTAAAGTGAAGCTCTTAAACAGTACGCAAGCTTAA 1546

```

Db      1796  TGAATCGATACCAAGAAATGCTAAAGTAGAGCACTCTTAACAGCTAGCAAGCTAA 1855
Oy      1547  CTTAGACTCAGAGATTTATAGCATCTCTGTGTAAGAGCTAAACTACTACAAACAATCT 1606
Db      1856  CTTAGACTCAGAGATTTATAGCATCTCTGTGTAAGAGCTAAACTACTACAAACAATCT 1915
Oy      1607  CGATGCTTTGGTATATAGGACTATACCTTAACCTGAAAAAGTAGAGATATACAGATGA 1666
Db      1916  CGATGCTTTGGTATATAGGACTATACCTTAACCTGAAAAAGTAGAGATATACAGATGA 1975
Oy      1667  CACCAACCGTATCATTAACCGTTATATGGCAAGCAGCCGGAAGAGAGATCTAGCTA 1726
Db      1976  CACCAACCGTATCATTAACCGTTATATGGCAAGCAGCCGGAAGAGAGATCTAGCTA 2035
Oy      1727  CCATTAGCTGGTGGCT 1742
Db      2036  TCATTAGCTTATGAT 2051

RESULT 4
SEDEXB  8931 bp      DNA      linear      BCT 18-APR-2005
LOCUS   S.equisimilis dexb, abc, lrp, skc, rel genes and ORF1.
DEFINITION
ACCESSION X72832
VERSION    X72832.1 GI:407876
KEYWORDS   abc gene; ABC transporter; dexb gene; dextran glucosidase; glucan
           1,6-alpha-glucosidase; leucine rich protein; LRP gene; ORF1; rel
           gene; skc gene; streptokinase; stringent response-like protein.
SOURCE
ORGANISM   Streptococcus dysgalactiae subsp. equisimilis
           Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
           Streptococcus.
REFERENCE 1 (bases 3621 to 6190)
AUTHORS   Malke,H., Roe,B. and Ferretti,J.J.
TITLE      Nucleotide sequence of the streptokinase gene from Streptococcus
           equisimilis H46A
JOURNAL   Gene 34 (2-3), 357-362 (1985)
PUBMED    2989113
REFERENCE 2 (bases 1 to 4188; 5790 to 8931)
AUTHORS   Mechohl,U., Steiner,K., Vettermann,S. and Malke,H.
TITLE      Genetic organization of the streptokinase region of the
           Streptococcus equisimilis H46A chromosome
JOURNAL   Mol.Gen. Genet. 241 (1-2), 129-140 (1993)
PUBMED    8232196
REFERENCE 3 (bases 1 to 8931)
AUTHORS   Malke,H.
TITLE      Direct Submission
JOURNAL   Submitted (05-MAY-1993) H. Malke, Institute for Molecular Biology,
           Jena University, Winzerlaer Str 10, 07708 Jena, FRG
COMMENT    Related sequences: K02986, M19346, X13399 & X13400.
FEATURES
         source
         1..8931
         /organism="Streptococcus dysgalactiae subsp. equisimilis"
         /mol_type="genomic DNA"
         /strain="H46A"
         /isolate="human group C strain"
         /sub_species="equisimilis"
         /db_xref="taxon:119602"
         /chromosome="streptokinase region"
         /clone_lib="E.coli plasmid library containing subfragments
         of the submitted seq:PSHD14/16, PSPV19, PSR2, PMF1,
         PCWP73, PRH10, PMX4"
         complement(89..1761)
         /gene="dexB"
         complement(89..115)
         /gene="dexB"
         complement(89..115)
         /note="hairpin loop"
         complement(136..1749)
         /gene="dexB"
         /BC_number="3.2.1.70"
         /codon_start=1
         /transl_table=11

```

```

/product="glucan 1,6-alpha-glucosidase"
/protein_id="CA51348.1"
/db_xref="GI:407877"
/db_xref="GOA:Q59905"
/db_xref="InterPro:IPR006047"
/db_xref="InterPro:IPR006589"
/db_xref="UniProt/Swiss-Prot:Q59905"
/translation="MOKOMHAKATVQIYPRSPKDTSGNGIGDLKGTISOLDYLOKLG
ITAIWISPVQSPMDNGNDYDISDYALAEVGRGMDMDLAAANERGKIKTIMDLVYN
HTSDERAHFVEARENPSPRDYIYRDEPNMIMSIPEGAMELDEASGOYIHLFSK
KQPDILWENAHVRQKLIYDMNFWIAGIGGPRMDVIDLICKIPDSEITGNGPRILHDL
KENNQATFNGHNDVMTVGETGATPEIARQYRPEENKELSMVFQEHVGLQHPKPAKPM
DYAEILDVPAIKTIIFSKMOTELKLGEMNSLFNNHDLPRVLSIWMGDSIYREGSALA
LAILHMRGTPYIYQGEETGMNYPKQDLTEVDLIESLVAKEAMENGVPAALVMS
IRKVGDMATPEMQSKDTHAGFSEAQETPLPKNPNQELINVADALANQDSIPPTYQO
LIARLKDQDMLEADYHLLPTADKVRAYQRFGEETIYIVNVSDQVRFKADLGALE
VVTITNDVDKVLTKHLQPMDAFCVKLSV"
complement(1757..1761)
/gene="dexB"
complement(1780..3051)
/gene="abc"
complement(1780..1810)
/gene="abc"
/note="hairpin loop"
complement(2482..2499)
/gene="abc"
/note="Walker motif B"
complement(2633..2965)
/gene="abc"
/codon_start=1
/transl_table=11
/product="ABC transporter"
/protein_id="CA51349.1"
/db_xref="GI:600072"
/db_xref="GOA:Q54086"
/db_xref="InterPro:IPR003439"
/db_xref="UniProt/TREMBL:Q54086"
/translation="MYEININHIYKYPNTTHYAVEDFDIDKKEIVVVGSGCKK
STLRMTAGLEDISBELKIDGSDVNDKSPKDIAMFQNYALYPMHTVYDMNAFGL
KLRKYKR"
complement(2644)
/gene="abc"
/note="(+) frame shift mutation in H46A"
complement(2828..2854)
/gene="abc"
/note="Walker motif A"
complement(2973..2976)
/gene="abc"
complement(3026..3031)
/gene="abc"
/note="alternative"
complement(3043..3048)
/gene="abc"
complement(3050..3051)
/gene="abc"
/note="TG motif"
complement(3244..4149)
/gene="lrp"
complement(3244..3261)
/gene="lrp"
/note="hairpin loop"
complement(3262..4110)
/gene="lrp"
/codon_start=1
/transl_table=11
/product="leucine rich protein"
/protein_id="CA51350.1"
/db_xref="GI:407878"
/db_xref="GOA:Q54087"
/db_xref="InterPro:IPR002197"
/db_xref="InterPro:IPR008931"
/db_xref="UniProt/Swiss-Prot:Q54087"
/translation="MELKDYFPEMQGPHFLGDKXWVSVEKGDQYVHFPKSCLSKKR

```

misc_feature	complement(3337. . 3423) /gene="ltp" /note="leucine zipper motif"
RBS	complement(4117. . 4120) /gene="ltp" complement(4141. . 4146) /gene="ltp" complement(4148. . 4149) /gene="ltp" /note="T5 motif"
-10_signal	4172. . 4368 /note="static DNA bending locus"
misc_feature	4392. . 5837 /gene="ekc"
gene	4392. . 4393 /gene="ekc" /note="T5 motif"
misc_feature	4395. . 4400 /gene="ekc"
-10_signal	4401. . 4406 /gene="ekc"
-10_signal	4428. . 4431 /gene="ekc"
RBS	4439. . 5761 /gene="ekc"
CDS	codon_start=1 transl_table=1 product="streptokinase" protein_id="CAA51351.1" db_xref="GI:407879" /db_xref="GOA:P00779" /db_xref="InterPro:IPR004093" /db_xref="InterPro:IPR008124" /db_xref="UniProt/Swiss-Prot:P00779" translation="MKNTLSPGMPALLPALPFGTVNSVOALSGPEWILLDRPSVNSQL VSVVAGTEGYNODISLKEPEIDLSRPAHGKTPQGLSPKSPAPADSAGMSKLEKE ADLAKIOLQILANYHSDNDYFEVIDPISADATITDRNKQVTFADKDSVTLPTPOVEE FLLSHAVRRPPEKPEPIONQAKSVDEYTVQFTPLNPDDPFGIKOTKLKTLAIDQ TITSELKAOQSLINKNHPGTYIYERDSSIVTHNDLFEFTLPMDEFTYRVNREG AYRIKSGSLGNEIINNTDLISEKYYLVKGEKPYDPFRSHLKLFTIKYVDVNEELL KSEOLLTSEBNLDPRDLYDPDRYAKLLYNNLDAPGIMDYTLTGKVDNHDNTRIT VYMGRRPGENASHTLADYDKRYTEREKEVSYLRTGTPIIPNPDK"
sig_peptide	4439. . 4516 /gene="ekc"
terminator	5796. . 5837 /gene="ekc"
CDS	complement(5859. . 6302) /note="unnamed protein product; ORF1" codon_start=1 transl_table=1 protein_id="CAA51352.1" /db_xref="GI:407880" /db_xref="GOA:O54088" /db_xref="InterPro:IPR003732" /db_xref="UniProt/Swiss-Prot:Q54088" translation="MKVLQRYKEASVSIIDKIGALINQGLLLVGVGPDPAABDLAY AVRKLVNRIISDADGKNKQNSIODIKGSIILSVSOFTLYADTKGNRPAPFGAAKPDMA SOFYRPFBOIADLPFVBRGVGADMQVSLINDGVTIILDTKCH"
conflict	6007. . 6009 citation=[1] replace="gc"
conflict	6162. . 6164 citation=[1] replace="cg"
RBS	complement(6310. . 6313) /note="ORF1"
gene	complement(6317. . 8575)

Query Match	Best Local Similarity	Score	DB 1	Length	8911
Matches	1150	Conservative	0	Mismatches	6
				Indels	0
				Gaps	0
587	TATTCGTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATGGTGT	646			
4516	TATTCGTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATGGTGT	4575			
647	TAGCGGTGCTGTACTGTGGAGGGAACCAATCAAGACATTAAGCTTTAAATTTTGAAT	706			
4576	TAGCGGTGCTGTACTGTGGAGGGAACCAATCAAGCTTTAAATTTTGAAT	46359			
707	CGATCTAACATACGACCGTCTCATAGAGGAAGAAGAGCAAGGCTTAAGTCCAAATC	766			
4636	CGATCTAACATACGACCGTCTCATAGAGGAAGAAGAGCAAGGCTTAAGTCCAAATC	46959			
767	AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGGAAAGCTACTACT	826			
4696	AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGGAAAGCTACTACT	47555			
827	AAAGCTATTCAAGAACCAATTGATGCTTACGTCACAGTAAACGACGACTTGAAGT	886			
4756	AAAGCTATTCAAGAACCAATTGATGCTTACGTCACAGTAAACGACGACTTGAAGT	48151			
887	CATTGATTTTGCAGCGATNGACACATTAATGGAACGCGACAGTCTATTGCTGCA	946			
4816	CATTGATTTTGCAGCGATNGACACATTAATGGAACGCGACAGTCTATTGCTGCA	48757			
947	CAAGATGCTTGGGTACCTTGCACCGACCACTTCCAGAAATTTTGTCTAAGCGACA	10060			
4876	CAAGATGCTTGGGTACCTTGCACCGACCACTTCCAGAAATTTTGTCTAAGCGACA	49355			
1007	TGTGCGCGTTAGACCATATATTAAGAAAAACAATTAACAAAAACCAAGGAAATCTGTGATGT	10666			
4936	TGTGCGCGTTAGACCATATATTAAGAAAAACAATTAACAAAAACCAAGGAAATCTGTGATGT	49959			
1067	GGAATATCTGACGATTACTCCCTTTAAACCTGATAGACGATTTCAACACAGGCTCTAA	11266			
4996	GGAATATCTGACGATTACTCCCTTTAAACCTGATAGACGATTTCAACACAGGCTCTAA	50555			
1127	AGATCTAAGCTATTGAAGAACTGATGCTGCTGACGATCAATCTCAAGAATTCT	11866			
5056	AGATCTAAGCTATTGAAGAACTGATGCTGCTGACGATCAATCTCAAGAATTCT	51151			
1187	AGCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTA	12466			
5116	AGCTCAAGCAACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTATGACGTA	51751			
1247	CTCTCATATCGTCACTCATGACATGACATTTCCGTAAGATTTTACCAATGATCAAGA	13066			
5176	CTCTCATATCGTCACTCATGACATGACATTTCCGTAAGATTTTACCAATGATCAAGA	52351			
1307	GTTTACTTACCGTGTAAATACCGGGAACAAGCTTATGAGTCAATATAAAATCTGCTCT	13666			
5236	GTTTACTTACCGTGTAAATACCGGGAACAAGCTTATGAGTCAATATAAAATCTGCTCT	52951			
1367	GAATGAAGAAATTAACAACACTGACCTGATCTCTGAGAAATTAAGCTTTAAAAAAG	14266			
5296	GAATGAAGAAATTAACAACACTGACCTGATCTCTGAGAAATTAAGCTTTAAAAAAG	53551			
1427	GGAAAAAGCGTATGATCCCTTTGATCGGACGCTTGAACCTGTTACCATCAATAGT	14866			
5356	GGAAAAAGCGTATGATCCCTTTGATCGGACGCTTGAACCTGTTACCATCAATAGT	54151			
1487	TGATGTGATACCAACGAATTTGCTTAAAGAGACGCTCTTAAACAGTACGGAACGTA	15466			
5416	TGATGTGATACCAACGAATTTGCTTAAAGAGACGCTCTTAAACAGTACGGAACGTA	54751			
1547	CTTGAAGCTTCAAGATTTATACGATCCCTGATTAAGGCTTAACTACTCTCAACAAATCT	16066			

Db 5476 CTAGAGCTTCAGAGATTATACGATCTCGTGATTAAGGCTAAACTACTTCAACAACATCT 5535
Qy 1607 CGATGCTTTGGATTATAGACTATACCTTAACCTGAGAAAGTAGAGATATACCGATGA 1666
Db 5536 CGATGCTTTGGATTATAGACTATACCTTAACCTGAGAAAGTAGAGATATACCGATGA 5595
Qy 1667 CACCAACCGTATCATTAACGTTTATATGAGCAAGCCGGAAGAGAGATGCTACTA 1726
Db 5596 CACCAACCGTATCATTAACGTTTATATGAGCAAGCCGGAAGAGAGATGCTACTA 5655
Qy 1727 CCATTTAGCTGTGTGT 1742
Db 5656 TCATTTAGCCTATGAT 5671
RESULT 5
ARI44000
LOCUS ARI44000 1242 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6210667.
ACCESSION ARI44000
VERSION ARI44000.1 GI:15105867
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1242)
Unclassified.
AUTHORS Reed,G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 5 03-APR-2001;
FEATURES
source 1..1242
/organism="Unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 54.6%; Score 1145.4; DB 6; Length 1242;
Best Local Similarity 99.5%; Pred. No.2.1e-297;
Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGGTGT 647
Db 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAAATGTTGT 60
Qy 648 ACCGTTGCTGCTACTGTGAGGGGAGAGATCAAGACATTAAGTCTTAATTTTGAATC 707
Db 61 ACCGTTGCTGCTACTGTGAGGGGAGAGATCAAGACATTAAGTCTTAATTTTGAATC 120
Qy 708 GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAAGTCCAAATCA 767
Db 121 GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGAGAGCTTAAGTCCAAATCA 180
Qy 768 AAACCATTTGCTACTGATAGTGGCGGAGATGTCATTAACCTTGAGAAAGCTGACTTACTA 827
Db 181 AAACCATTTGCTACTGATAGTGGCGGAGATGTCATTAACCTTGAGAAAGCTGACTTACTA 240
Qy 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACAGACGCTACTTTAGAGTGC 887
Db 241 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACAGACGCTACTTTAGAGTGC 300
Qy 888 ATTGATTTTGAAGGAGATGAACCATTAAGTCAAGAGGCGAAGGCTTACTTTGCGAC 947
Db 301 ATTGATTTTGAAGGAGATGAACCATTAAGTCAAGAGGCGAAGGCTTACTTTGCGAC 360
Qy 948 AAAGATGTTGGTGAACCTTGGCCAGCCAACTGTCCAAAGATTTTGTCTAAGCGACAT 1007
Db 361 AAAGATGTTGGTGAACCTTGGCCAGCCAACTGTCCAAAGATTTTGTCTAAGCGACAT 420
Qy 1008 GTGCGCGTTAGAACCATATATAAGAAAAACAATACAAAGCGAAATGTTGATGATG 1067
Db 421 GTGCGCGTTAGAACCATATATAAGAAAAACAATACAAAGCGAAATGTTGATGATG 480
Qy 1068 GAATATACTGTACAGTTTACTCCCTTAAACCCGTGATGACGATTTGACAGCAGGCTCAAA 1127

Db 481 GAATATACTGTACAGTTTACTCCCTTAAACCCGTGATGACGATTTGACAGCAGGCTCAAA 540
Qy 1128 GATCTAACATTTGAAAAACATGATATGCTGATGACACCATCATCTCAAGAAATTACTA 1187
Db 541 GATCTAACATTTGAAAAACATGATATGCTGATGACACCATCATCTCAAGAAATTACTA 600
Qy 1188 GCTCAAGCAACAAGCATTTTAAACAAAAACCAAGGCTATTAAGATTTATGAAGTAC 1247
Db 601 GCTCAAGCAACAAGCATTTTAAACAAAAACCAAGGCTATTAAGATTTATGAAGTAC 660
Qy 1248 TCCTCAATCGTCACTGATGACATGACATTTCCGTACGATTTTAAACATGATCAAGAG 1307
Db 661 TCCTCAATCGTCACTGATGACATGACATTTCCGTACGATTTTAAACATGATCAAGAG 720
Qy 1308 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCATTAATAAATCTGTCTG 1367
Db 721 TTTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCATTAATAAATCTGTCTG 780
Qy 1368 AATGAGAAATTAACCAACTGACCTGATCTGTGAGAAATATTAAGTCTTAATAAAGG 1427
Db 781 AATGAGAAATTAACCAACTGACCTGATCTGTGAGAAATATTAAGTCTTAATAAAGG 840
Qy 1428 GAAAAGCGATATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACCATCAATACGTT 1487
Db 841 GAAAAGCGATATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACCATCAATACGTT 900
Qy 1488 GATGTCATTCACCAAGAAATGCTTAAAAAGTAGAGAGCTTTAAACGCTAGCGAAGCTTAC 1547
Db 901 GATGTCATTCACCAAGAAATGCTTAAAAAGTAGAGAGCTTTAAACGCTAGCGAAGCTTAC 960
Qy 1548 TTAGACTTCAGAGATTATATGATCTCGTGTATAGGCTTAACCTTACCAACAATCTC 1607
Db 961 TTAGACTTCAGAGATTATATGATCTCGTGTATAGGCTTAACCTTACCAACAATCTC 1020
Qy 1608 GATCTTTTGGTATTAATGACCTATACCTTAACCTGAAAAAGTAGAGATTAACAGATGAC 1667
Db 1021 GATCTTTTGGTATTAATGACCTATACCTTAACCTGAAAAAGTAGAGATTAACAGATGAC 1080
Qy 1668 ACCAACCGTATCATTAACCGTTTATATGAGGCAAGCGACCCGGAAGAGAGATGCTACTAC 1727
Db 1081 ACCAACCGTATCATTAACCGTTTATATGAGGCAAGCGACCCGGAAGAGAGATGCTACTAC 1140
Qy 1728 CATTTAGCTGTGT 1742
Db 1141 CATTTAGCTGTATGAT 1155
RESULT 6
E00522 2568 bp DNA linear PAT 29-SEP-1997
LOCUS E00522
DEFINITION DNA fragment of plasmid pMF1 into which DNA encoding streptokinase is inserted.
ACCESSION E00522
VERSION E00522.1 GI:2168801
KEYWORDS JP 1985237995-A/1.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 2568)
AUTHORS Jiyosefu,J.P. and Hornuto,M.
TITLE RECOMBINATION VECTOR FOR PRODUCING STREPTOKINASE
JOURNAL Patent: JP 1985237995-A 1 26-NOV-1985;
PHILIPS PETROLEUM CO
OS Streptococcus equisimilis
PN JP 1985237995-A/1
PD 26-NOV-1985
PF 09-OCT-1984 JP 1984212403
PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PI
PJ JIYOSSEFU JIJI FUJERETSUJI, HORUSUTO MARUKI
PC C12N15/00,C12N1/20,C12N9/70,(C12N15/00,C12R1.46),(C12N1/20,PC C12R1.19),

```

PC (C12N9/70,C12R1.19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: Strain=Streptococcus equismilis H46A; CC *source:
library=Streptococcus equismilis H46 library; CC *source:
clone=Lambda 147 skc clone;
FH Key Location/Qualifiers
FT 5'UTR 1..818
FT sig_peptide 819..896
FT /product='streptokinase signal peptide' FT
mat_peptide 897..2138
FT /product='streptokinase'
FT CDS 819..2143
FT /product='streptokinase'
FT 3'UTR 2142..2569.
Location/Qualifiers
1..2568
/organism='Streptococcus dysgalactiae subsp. equismilis'
/mol_type='genomic DNA'
/db_xref='taxon:119602'

ORIGIN
Query Match 54.5%; Score 1141.6; DB 6; Length 2568;
Best Local Similarity 99.1%; Pred. No. 2.3e-295;
Matches 1147; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 587 TATGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGTTGT 646
DB 896 TATGTGTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATTTGTTGT 955
QY 647 TAGCGTGTGCTGCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAAT 706
DB 956 TAGCGTGTGCTGCTGTGAGGGAGCAATCAAGACATTAGTCTTAATTTTGAAT 1015
QY 707 CGATCTAACATACGACGCTCATGGAGGAAGACAGCAAGGCTTAGTCCAAATC 766
DB 1016 CGATCTAACATACGACGCTCATGGAGGAAGACAGCAAGGCTTAGTCCAAATC 1075
QY 767 AAAACATTGTGCTAGTAGTGGCCGAGTGCATTAACCTTAGAAGCTGACTTACT 826
DB 1076 AAAACATTGTGCTAGTAGTGGCCGAGTGCATTAACCTTAGAAGCTGACTTACT 1135
QY 827 AAAAGCTATTCAAGAACATTTGCTTAACTTCAAGTAAAGCACTACTTTGAGT 886
DB 1136 AAAAGCTATTCAAGAACATTTGCTTAACTTCAAGTAAAGCACTACTTTGAGT 1195
QY 887 CATTGATTTTTCAGAGGATGCAACATTAATTCGAAAGGCAAGTCTACTTTGCTGA 946
DB 1196 CATTGATTTTTCAGAGGATGCAACATTAATTCGAAAGGCAAGTCTACTTTGCTGA 1255
QY 947 CAAAGTGGTTCGTTAAGCTTCCGACCCCAACCTGTCCAGAAATTTTGTAAAGCGACA 1006
DB 1256 CAAAGTGGTTCGTTAAGCTTCCGACCCCAACCTGTCCAGAAATTTTGTAAAGCGACA 1315
QY 1007 TGTGGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAATCTGTGATGT 1066
DB 1316 TGTGGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAATCTGTGATGT 1375
QY 1067 GGAATATATCTGTAAGTTTACTCCCTTAAACCTGTATGAGATTTTGAACAGGCTTCAA 1126
DB 1376 GGAATATATCTGTAAGTTTACTCCCTTAAACCTGTATGAGATTTTGAACAGGCTTCAA 1435
QY 1127 AGATCTAGATGTAATTAAGAAAACTAGCTATCGGTGACACATGACATCTTAAAGTAACT 1186
DB 1436 AGATCTAGATGTAATTAAGAAAACTAGCTATCGGTGACACATGACATCTTAAAGTAACT 1495
QY 1187 AGCTCAAGCAAAAGATTTTAAACAAACCAAGGCTATACGATTTATGAAAGCTGA 1246
DB 1496 AGCTCAAGCAAAAGATTTTAAACAAACCAAGGCTATACGATTTATGAAAGCTGA 1555

```

```

QY 1247 CTCCTCAATGCTACCTACATGACAAATGATTTTCGGTACGATTTTACCAATGATCAAGA 1306
DB 1556 CTCCTCAATGCTACCTACATGACAAATGATTTTCGGTACGATTTTACCAATGATCAAGA 1615
QY 1307 GTTACTTACCTGTTTAAATTCGGAAACAAAGCTTATGATCAATTAATAATCTGTCT 1366
DB 1616 GTTACTTACCTGTTTAAATTCGGAAACAAAGCTTATGATCAATTAATAATCTGTCT 1675
QY 1367 GAATGAAGAAATTAACAACTGACCTGATCTCGAAGAAATTAATACGCTTAAATAAAG 1426
DB 1676 GAATGAAGAAATTAACAACTGACCTGATCTCGAAGAAATTAATACGCTTAAATAAAG 1735
QY 1427 GGAAGACCGTATGATCCCTTTGATTCGAGTCACTTGAAGCTTCAACATCAATACGT 1486
DB 1736 GGAAGACCGTATGATCCCTTTGATTCGAGTCACTTGAAGCTTCAACATCAATACGT 1795
QY 1487 TGATGTGATACCAAGCAATGCTTAAAGTGAAGAGCTTTAAACGCTAAGCAAGTAA 1546
DB 1796 TGATGTGATACCAAGCAATGCTTAAAGTGAAGAGCTTTAAACGCTAAGCAAGTAA 1855
QY 1547 CTTAGACTTCAGAGATTTATACGATCCCTCGATTAAGGCTAACTACTTCAACAATCT 1606
DB 1856 CTTAGACTTCAGAGATTTATACGATCCCTCGATTAAGGCTAACTACTTCAACAATCT 1915
QY 1607 CGATGCTTTTGGTATTAATGACATTAACCTTAAGTGAAGAAATGAGATAATCAAGATGA 1666
DB 1916 CGATGCTTTTGGTATTAATGACATTAACCTTAAGTGAAGAAATGAGATAATCAAGATGA 1975
QY 1667 CACCAACCGTATCAATTAACGTTTATATGGGAAGGACCCGAAGAGAAATGCTAGCTA 1726
DB 1976 CACCAACCGTATCAATTAACGTTTATATGGGAAGGACCCGAAGAGAAATGCTAGCTA 2035
QY 1727 CCATTAGCTGTGTGT 1742
DB 2036 CCATTAGCTGTGTAT 2051

RESULT 7
LOCUS 105204 1401 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 5 from Patent EP 0248227.
ACCESSION 105204
VERSION 105204.1 GI:591209
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1401)
AUTHORS
Hagenson, M.J. and Stroman, D.W.
TITL
Yeast production of streptokinase
JOURNAL
Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES
source
1..1401
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 54.3%; Score 1139; DB 6; Length 1401;
Best Local Similarity 99.1%; Pred. No. 1.1e-295;
Matches 1145; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCAATTTGTTGT 647
DB 10 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCAATTTGTTGT 69
QY 648 AGCTTCTGCTGATCTGTGAGGGGAGCAATTAAGACATTAATCTTAAATTTTGAATC 707
DB 70 AGCTTCTGCTGATCTGTGAGGGGAGCAATTAAGACATTAATCTTAAATTTTGAATC 129
QY 708 GATCTAACATCAAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAAATCA 767
DB 130 GATCTAACATCAAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAAATCA 189

```


QY 768 AAACATTTGCTACTGATAGTGGCGGATGTCACTATAAATTGAGAAAGCTGACTTA 827
DB 190 AAACATTTGCTACTGATAGTGGCGGATGTCACTATAAATTGAGAAAGCTGACTTA 249
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCCAGTACGACGACGACGACGACG 887
DB 250 AAGGCTATTCAAGAACATTTGATCGCTAACGTCCAGTACGACGACGACGACGACG 309
QY 888 ATTGATTTTGGAAAGGATGCAACCATTAAGTACGATGAAAGGCGAAGGCTACTTTGCTGAC 947
DB 310 ATTGATTTTGGAAAGGATGCAACCATTAAGTACGATGAAAGGCGAAGGCTACTTTGCTGAC 369
QY 948 AAAAGATGTTCCGCTAACCTTGGCGAACCCAACTGTCCAGAAATTTTGGCTAACGCGACAT 1007
DB 370 AAAAGATGTTCCGCTAACCTTGGCGAACCCAACTGTCCAGAAATTTTGGCTAACGCGACAT 429
QY 1008 GTGGCGGTTGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1067
DB 430 GTGGCGGTTGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 489
QY 1068 GAATATCTAGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACCGAGTCTCAA 1127
DB 490 GAATATCTAGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACCGAGTCTCAA 549
QY 1128 GATATCTAGTACATTTGAAAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 1187
DB 550 GATATCTAGTACATTTGAAAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 609
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCGAGCTATAAGATTTATGAAAGCTGAC 1247
DB 610 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCGAGCTATAAGATTTATGAAAGCTGAC 669
QY 1248 TCCTCAATGCTACATCACTAGCACTGACATTTTCCGATCCATTTTACCAATGATCAAGAG 1307
DB 670 TCCTCAATGCTACATCACTAGCACTGACATTTTCCGATCCATTTTACCAATGATCAAGAG 729
QY 1308 TTTACTTACCGTGTAAAAAATCGGGAACAGCTTATAGATCAATTAATAAATCGTGTG 1367
DB 730 TTTACTTACCGTGTAAAAAATCGGGAACAGCTTATAGATCAATTAATAAATCGTGTG 789
QY 1368 AATGAAAGAAATTAACAACAACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGG 1427
DB 790 AATGAAAGAAATTAACAACAACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGG 849
QY 1428 GAAAAGCCGTATGATCCCTTTGATCGGATCACTTGAACCTGTCACATCAATACGTT 1487
DB 850 GAAAAGCCGTATGATCCCTTTGATCGGATCACTTGAACCTGTCACATCAATACGTT 909
QY 1488 GATGTCGATACCAAGAAATTTGCTAAAGAGGACGCTCTTAACGCTGCGAAGCGTAAC 1547
DB 910 GATGTCGATACCAAGAAATTTGCTAAAGAGGACGCTCTTAACGCTGCGAAGCGTAAC 969
QY 1548 TTGAGCTTCAGAGATTTTATACGATCTCTGATTAAGGCTTAACTACTCTCAACAATCTC 1607
DB 970 TTGAGCTTCAGAGATTTTATACGATCTCTGATTAAGGCTTAACTACTCTCAACAATCTC 1029
QY 1608 GATGCTTTTGTATTAATGACATTAACCTTAACGTAAGAAAGTAAAGATTAATCAAGTAC 1667
DB 1030 GATGCTTTTGTATTAATGACATTAACCTTAACGTAAGAAAGTAAAGATTAATCAAGTAC 1089
QY 1668 ACCAACCGTATACATACCGTTTATATGCGCAAGCGACCGCAAGAGAAATGCTACTAC 1727
DB 1090 ACCAACCGTATACATACCGTTTATATGCGCAAGCGACCGCAAGAGAAATGCTACTAC 1149
QY 1728 CATTTAGCTGTGT 1742
DB 1150 CATTTAGCTGTGT 1164

RESULT 8
AR068768
LOCUS

DEFINITION Sequence 19 from patent US 5854049. 2566 bp DNA linear PAT 29-SEP-1999

ACCESSION AR068768
VERSION AR068768.1 GI:6000975
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2566)
TITLE Reel, G.L.
JOURNAL Plasmid-resistant streptokinase
PATENT Patent: US 5854049-A 19 29-DEC-1998;
FEATURES Location/Qualifiers
Source 1..2566
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 54.0%; Score 1132.4; DB 6; Length 2566;
Best Local Similarity 99.3%; Pred. No. 7.1e-294;
Matches 1148; Conservative 0; Mismatches 6; Indels 2; Gaps 1;
QY 587 TATTGCTGACCTGAGTGGCTGCTGACCCGTCATCTGTCAACAAACAGCCAAATGGTTGT 646
DB 896 TATTGCTGACCTGAGTGGCTGCTGACCCGTCATCTGTCAACAAACAGCCAAATGGTTGT 955
QY 647 TAGGTTGCTGCTGATGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 706
DB 956 TAGGTTGCTGCTGATGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 1015
QY 707 GATCTAACATCAAGACCTGCTCAATGAGGAAAGACAGAGCAAGCTTAAGTCAAAATC 766
DB 1016 GATCTAACATCAAGACCTGCTCAT - AGGAAAGACAGAGCAAGCTTAAGTCAAAATC 1073
QY 767 AAAACATTTGCTACTGATATGTTGGCGGATGTCACTATAAATTGAGAAAGCTGACT 826
DB 1074 AAAACATTTGCTACTGATATGTTGGCGGATGTCACTATAAATTGAGAAAGCTGACT 1133
QY 827 AAAAGCTATTCAGAAATGATGCTAAGCTCAAGTACAGACGACTACTTGAAGT 886
DB 1134 AAAAGCTATTCAGAAATGATGCTAAGCTCAAGTACAGACGACTACTTGAAGT 1193
QY 887 CATTGATTTTGAAGCGATCAACCATTAAGTCAAGACGCAAGGCTTACTTTGCTGA 946
DB 1194 CATTGATTTTGAAGCGATCAACCATTAAGTCAAGACGCAAGGCTTACTTTGCTGA 1253
QY 947 CAAGATGTTGCGTAACTTTGCCGACCCAACTGTCTCAAGAAATTTTGTACGCGACA 1006
DB 1254 CAAGATGTTGCGTAACTTTGCCGACCCAACTGTCTCAAGAAATTTTGTACGCGACA 1313
QY 1007 TGTGCGGCTTAAGACCATATAAGAAACCAATTCAGAAACCAAGCGAAATCTGTGATGT 1066
DB 1314 TGTGCGGCTTAAGACCATATAAGAAACCAATTCAGAAACCAAGCGAAATCTGTGATGT 1373
QY 1067 GGAATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTGAGACGAGTCTCAA 1126
DB 1374 GGAATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTGAGACGAGTCTCAA 1433
QY 1127 AGATCTAAGCTATTTGAAAAACCTAGCTATCGGTGACACCATCACTCAAGAAATTA 1186
DB 1434 AGATCTAAGCTATTTGAAAAACCTAGCTATCGGTGACACCATCACTCAAGAAATTA 1493
QY 1187 AGCTCAAGGCAAAAGCATTTTAAACCAAAACCAACCGGCTATACGATTTATGAAAGTGA 1246
DB 1494 AGCTCAAGGCAAAAGCATTTTAAACCAAAACCAACCGGCTATACGATTTATGAAAGTGA 1553
QY 1247 CTCCTCAATGCTCACTGATGACATGATTTTCCGTAGCATTTTACCAATGATGATCAAGA 1306
DB 1554 CTCCTCAATGCTCACTGATGACATGATTTTCCGTAGCATTTTACCAATGATGATCAAGA 1613
QY 1307 GTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATATAAATCTGTGCT 1366
DB 1614 GTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATATAAATCTGTGCT 1673
QY 1367 GAATGAAGAAATTAACAACAACCTGATCTCTGAGAAATATTACGCTCTTAAAAAAGG 1426

Db 1674 GATATAGAAATTAACACACGACCTGATCTCTGAGAAATATAGTCTTAAAGG 1733
Qy 1427 GAAAAGCCGATGATCCCTTTGATCGACGACCTTGAACCTGTCACATCAATCGT 1486
Db 1734 GAAAAGCCGATGATCCCTTTGATCGACGACCTTGAACCTGTCACATCAATCGT 1793
Qy 1487 TGATGTCATACCAAGAAATGTCGTAAGAGAGAGCTTTAAGAGTGAAGCAAGTAA 1546
Db 1794 TGATGTCATACCAAGAAATGTCGTAAGAGAGAGCTTTAAGAGTGAAGCAAGTAA 1853
Qy 1547 CTTAAGCTTGAAGATTTAAGATCTGTCGTAAGAGTGAAGCAAGTAACTTCAACATCT 1606
Db 1854 CTTAAGCTTGAAGATTTAAGATCTGTCGTAAGAGTGAAGCAAGTAACTTCAACATCT 1913
Qy 1607 CGATCTTTGATGATGACCTAATCTTAACTGGAAGTGAAGTAAATCAAGTGA 1666
Db 1914 CGATCTTTGATGATGACCTAATCTTAACTGGAAGTGAAGTAAATCAAGTGA 1973
Qy 1667 CACCAACCGTATCATACCGTTTATATGCGCAAGCCGAGAGAGATCTAGCTA 1726
Db 1974 CACCAACCGTATCATACCGTTTATATGCGCAAGCCGAGAGAGATCTAGCTA 2033
Qy 1727 CCATTTAGCTGCTGCT 1742
Db 2034 TCATTTAGCTATGAT 2049

RESULT 9

A20027 1458 bp DNA linear PAT 14-JUL-1995
LOCUS
DEFINITION SEQ ID NO: 35; Synthetic nucleotide sequence for Hsrudin-streptokinase fusion protein.
ACCESSION A20027
VERSION A20027.1 GI:1247862
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1458)
AUTHORS
TITLE PROTEINS AND NUCLEIC ACIDS
JOURNAL Patent: WO 9109125-A 35 27-JUN-1991;
FEATURES Location/Qualifiers
source 1..1458
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 53.7%; Score 1125.4; DB 6; Length 1458;
Best Local Similarity 97.8%; Pred. No. 5.2e-292;
Matches 1141; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 576 ACCGATGTTGATGTCGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACACAGC 635
Db 196 ATCGAAGTAGAATGCTGACCTAGTGGCTGCTAGACCGTCCATCTGTCACACAGC 255
Qy 636 CAATTGTTGTTAGCGTGTGCTGCTGCTGTTGAGGGGACGAATCAAGACATTTAGTCTTAA 695
Db 256 CAATTGTTGTTAGCGTGTGCTGCTGCTGTTGAGGGGACGAATCAAGACATTTAGTCTTAA 315
Qy 696 TTTTGTGAATGATCTTAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTA 755
Db 316 TTTTGTGAATGATCTTAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTA 375
Qy 756 AGTCCAAATGAAAACCTTTGCTACTGATAGTGGCGCGATGTCATATACTTGAAGAA 815
Db 376 AGTCCAAATGAAAACCTTTGCTACTGATAGTGGCGCGATGTCATATACTTGAAGAA 435
Qy 816 GCTGACTTACTTAAGGCTATTCAAGAACTATGATGCTTAAGCTCAGAGTAAAGAGAC 875
Db 436 GCTGACTTACTTAAGGCTATTCAAGAACTATGATGCTTAAGCTCAGAGTAAAGAGAC 495

Qy 876 TACTTGAAGTCAATTGATTTTGGCAAGCATCAACATTTACTGATCGAAAGGCAAGGTC 935
Db 496 TACTTGAAGTCAATTGATTTTGGCAAGCATCAACATTTACTGATCGAAAGGCAAGGTC 555
Qy 936 TACTTGTGTCGAAAGATGTTGCGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTG 995
Db 556 TACTTGTGTCGAAAGATGTTGCGTAACTTTGCCGACCAACCTGTCCAAAGAAATTTTG 615
Qy 996 CTAAAGCGACATGAGCGGTGAGACCATATTAAGAAAACCAATTAACCAAGCCGAA 1055
Db 616 CTAAAGCGACATGAGCGGTGAGACCATATTAAGAAAACCAATTAACCAAGCCGAA 675
Qy 1056 TCTGTTGATGGAATTAATCTGACAGTTTATCTCCCTTAAACCTGATGACGATTTGAGA 1115
Db 676 TCTGTTGATGGAATTAATCTGACAGTTTATCTCCCTTAAACCTGATGACGATTTGAGA 735
Qy 1116 CCAGTCTCAAGATTAAGCTTAATGAAAACCTAGCTATCGGTGACACATCAATCT 1175
Db 736 CCAGTCTCAAGATTAAGCTTAATGAAAACCTAGCTATCGGTGACACATCAATCT 795
Qy 1176 CAAGAAATTAAGTCAAGACCAAGCAATTTTAAACAAAACCAACCGGCTATTCGATT 1235
Db 796 CAAGAAATTAAGTCAAGACCAAGCAATTTTAAACAAAACCAACCGGCTATTCGATT 855
Qy 1236 TATGAAGTGAATCTCTCAATGTCATCTCATGACATGACATTTTCCGTAGCAATTTTACCA 1295
Db 856 TATGAAGTGAATCTCTCAATGTCATCTCATGACATGACATTTTCCGTAGCAATTTTACCA 915
Qy 1296 ATGATCAAGAGTTTACTTACCGTGTAAATCGGAAACAAAGCTTATAGATCAATTA 1355
Db 916 ATGATCAAGAGTTTACTTACCGTGTAAATCGGAAACAAAGCTTATAGATCAATTA 975
Qy 1356 AAATCTGCTGATGATGAAGAAATTAACAACAACCTGATCTGTGAGAAATTTACGTC 1415
Db 976 AAATCTGCTGATGATGAAGAAATTAACAACAACCTGATCTGTGAGAAATTTACGTC 1035
Qy 1416 CTTAAAAAAGGGGAAAGCCGTATGATCCCTTATGCGAGTCACCTGAAACCTTGACC 1475
Db 1036 CTTAAAAAAGGGGAAAGCCGTATGATCCCTTATGCGAGTCACCTGAAACCTTGACC 1095
Qy 1476 ATCAATACGTTGATGTCATACCAAGAAATGCTTAAAGTGAAGAGCTTTAAAGCT 1535
Db 1096 ATCAATACGTTGATGTCATACCAAGAAATGCTTAAAGTGAAGAGCTTTAAAGCT 1155
Qy 1536 AGCGAAGCTAATCTTGAATTTATGAGATTTTATGAGATCTCTGCTGAATAGGCTTAACTACTC 1595
Db 1156 AGCGAAGCTAATCTTGAATTTATGAGATTTTATGAGATCTCTGCTGAATAGGCTTAACTACTC 1215
Qy 1596 TACAACAATCTCGATGCTTTTGTATATGAGATTTTATGAGATCTTAACTGAAAAATGAGAT 1655
Db 1216 TACAACAATCTCGATGCTTTTGTATATGAGATTTTATGAGATCTTAACTGAAAAATGAGAT 1275
Qy 1656 AATCAGATGACCAACCGATATCAATACGTTTATATGAGGCAAGCCGAAAGAGAG 1715
Db 1276 AATCAGATGACCAACCGATATCAATACGTTTATATGAGGCAAGCCGAAAGAGAG 1335
Qy 1716 AATGCTAGCTTACATTTAGCTGTGCT 1742
Db 1336 AATGCTAGCTTACATTTAGCTTATGAT 1362

RESULT 10
113215 1458 bp DNA linear PAT 26-JUL-1995
LOCUS
DEFINITION Sequence 42 from patent US 5434073.
ACCESSION 113215
VERSION 113215.1 GI:910563
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1458)

AUTHORS Dawson, K., Hunter, M.G. and Czaplewski, L.G.
 TITLE Fibrinolytic and anti-thrombotic cleavable dimers
 JOURNAL Patent: US 5434073-A 42 18-JUL-1995;
 FEATURES Location/Qualifiers
 source 1..1458
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 53.7%; Score 1125.4; DB 6; Length 1458;
 Best Local Similarity 97.8%; Pred. No. 5.2e-292;
 Matches 1141; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

576 ACCGATGTCGATTCCTGAGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAAGC 635
 196 ATCGAAGTGAAGATTGCTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACAAGC 255
 636 CAATTGTTGTTAGCGTTCCTGCTGCTGTTGAGGGGACGAATCAAGACATTAGCTTTAA 695
 256 CAATTAGTTGTTAGCGTTCCTGCTGCTGTTGAGGGGACGAATCAAGACATTAGCTTTAA 315
 696 TTTTGTGAATGATTCATACGACCTGCTCATGAGGAAAGACAGGCAAGGCTTA 755
 316 TTTTGTGAATGATTCATACGACCTGCTCATGAGGAAAGACAGGCAAGGCTTA 375
 756 AGTCCAAATCAAAACCATTTGCTATGATGAGGCGGCTGCTCATATACTTGAAGAA 815
 376 AGTCCAAATCAAAACCATTTGCTATGATGAGGCGGCTGCTCATATACTTGAAGAA 435
 816 GCTGACTTACTAAAGGCTATTCAGAAACATTTAGCTTACGCTCAAGTAAAGACGAC 875
 436 GCTGACTTACTAAAGGCTATTCAGAAACATTTAGCTTACGCTCAAGTAAAGACGAC 495
 876 TACTTTGAGCTCATTTGTTGCAACCGTACCATTTATCTGATGGAACGGCAAGGTC 935
 496 TACTTTGAGCTCATTTGTTGCAACCGTACCATTTATCTGATGGAACGGCAAGGTC 555
 936 TACTTTGAGCTCATTTGTTGCAACCGTACCGTACCGACCACTGCTCAAGAAATTTTG 995
 556 TACTTTGAGCTCATTTGTTGCAACCGTACCGTACCGACCACTGCTCAAGAAATTTTG 615
 996 CTAAACCGGACATGTGGCGCTTAGACCATATTAAGAAACCAATACAAACCAAGCGAA 1055
 616 CTAAACCGGACATGTGGCGCTTAGACCATATTAAGAAACCAATACAAACCAAGCGAA 675
 1056 TCTGTTGATGTGGAATATCTGTACAGTTTACTCCTTAAACCTGATGACGATTTTCA 1115
 676 TCTGTTGATGTGGAATATCTGTACAGTTTACTCCTTAAACCTGATGACGATTTTCA 735
 1116 CCAAGTCTCAAGATCTAAGCTATGAAACCATAGCTATGCGTACACCATCACTCT 1175
 736 CCAAGTCTCAAGATCTAAGCTATGAAACCATAGCTATGCGTACACCATCACTCT 795
 1176 CCAAGATTAATCTAGCTCAAGCAACAACATTTTAAACAAACCAACCAAGCTTACGAT 1235
 796 CCAAGATTAATCTAGCTCAAGCAACAACATTTTAAACCAACCAAGCTTACGAT 855
 1236 TATGACGCTGACTCTCTGAATCTGCTACTGATGACATGACATTTTCCGATGATTTTCA 1295
 856 TATGACGCTGACTCTCTGAATCTGCTACTGATGACATGACATTTTCCGATGATTTTCA 915
 1296 ATGATCAAGATTTTACTTACCGTGTAAACCAACCGGAAACAGCTTATAGATCAATTA 1355
 916 ATGATCAAGATTTTACTTACCGTGTAAACCAACCGGAAACAGCTTATAGATCAATTA 975
 1356 AAATCTGCTCTGAATGAGAAATTAACAACATGACCTGATCTCTGAGAAATTTTACGTC 1415
 976 AAATCTGCTCTGAATGAGAAATTAACAACATGACCTGATCTCTGAGAAATTTTACGTC 1035
 1416 CTTAAACCAAGGGAAGAGCGGATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAAC 1475
 1036 CTTAAACCAAGGGAAGAGCGGATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAAC 1095

Qy 1476 ATCAATAGCTGATGATGCTGATACCAACGAATTTGCTAAAAAGTAGCAGCTCTTAACAGCT 1535
 Db 1096 ATCAATAGCTGATGATGCTGATACCAACGAATTTGCTAAAAAGTAGCAGCTCTTAACAGCT 1155
 Qy 1536 AGCGAAGTAACTTAAGCTTCAAGATTTTATACATCTCTGATTAAGGCTTAACCTACTC 1595
 Db 1156 AGCGAAGTAACTTAAGCTTCAAGATTTTATACATCTCTGATTAAGGCTTAACCTACTC 1215
 Qy 1596 TACAAACATCTGATGCTTTTGGTATTAAGGACCTATACCTTAACCTGAAAAAGTAGAGAT 1655
 Db 1216 TACAAACATCTGATGCTTTTGGTATTAAGGACCTATACCTTAACCTGAAAAAGTAGAGAT 1275
 Qy 1656 AATCAGATGACACCAACCGTATCATACCGTTATATAGGCAAGCAGCCGAAGAGAG 1715
 Db 1276 AATCAGATGACACCAACCGTATCATACCGTTATATAGGCAAGCAGCCGAAGAGAG 1335
 Qy 1716 AATGCTAGCTACCATTTAGCTGCTGCT 1742
 Db 1336 AATGCTAGCTACCATTTAGCTGCTGCTATGAT 1362

RESULT 11

S46536 1245 bp DNA linear BCT 08-MAY-1993
 LOCUS SKC-2=streptokinase [Streptococcus equisimilis, group C, ATCC 9542,
 Genomic, 1245 nt].
 S46536
 VERSION S46536.1 GI:257196
 KEYWORDS
 SOURCE Streptococcus dysgalactiae subsp. equisimilis
 ORGANISM Streptococcus dysgalactiae subsp. equisimilis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.

REFERENCE
 1 (bases 1 to 1245)
 Estrada, M.P., Hernandez, L., Perez, A., Rodriguez, P., Serrano, R.,
 Rubiera, R., Pedraza, A., Padron, G., Antuch, W., de la Fuente, J., et
 al.

TITLE High level expression of streptokinase in *Escherichia coli*

JOURNAL Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)

PUBMED 1368792

REMARK Genbank staff at the National Library of Medicine created this
 entry [NCBI gisbseq 115306] from the original journal article.

FEATURES

source

Location/Qualifiers

1..1245

/organism="Streptococcus dysgalactiae subsp. equisimilis"

/mol_type="genomic DNA"

/db_xref="taxon:119602"

/notes="group: C"

1..1245

/gene="SKC-2"

1..1245

/gene="SKC-2"

/note="streptokinase; Mature protein lacks initial Met"

/codon_start=1

/product="SKC-2"

/protein_id="AAC60418.1"

/db_xref="GI:257196"

/translation="IAGPEWLIDRPSVNSQLVSVAGTEGNODISLKEFDLTS
 RPAHGKTEBOGLSPKSPFATDSGAMHKEKDLALAIOLIANHSNDYEVD
 FADATITDNGKGVYFADKDGSTLPLPOVPEFLSGHVRVPYKEXPIONAKSVY
 EYTVOPFLPLNDPDDFRPGIDKTLKTLAIGDITTSOELALAOISINKTHGVYIE
 RDSISYTHNDITRTILPMQDEFTYHAKNEQAVEINKSGSINEINNTLISKYV
 LKRGKPYDPRDSHLFTIKYVDVNTNELSBOULTASERNLDFREDLDPDKAK
 LLYNNLDAFGIMDYTLTGKVEDHNDITNRIITYVMGRPRENASYHLAADKQRYEE
 EREVSYLRATGTPIDNPNDX"

ORIGIN

Query Match 53.7%; Score 1124.6; DB 1; Length 1245;
 Best Local Similarity 98.4%; Pred. No. 8.5e-292;
 Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATGCTGTTGTT 647

Db 1 ATTGCTGAGACTGAGTGGCTGCTGAGACCGCTCATCTGCAACACAGCCAAATTAGTGT 60
Qy 648 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTAAATC 707
Db 61 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGTAAAT 120
Qy 708 GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
Db 121 GACCTAACATCAAGACCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 180
Qy 768 AAACATTGCTGATGATGAGGGGCGATGTCACATAAATTGAGAAAGCTGACTTACTA 827
Db 181 AAACATTGCTGATGATGAGGGGCGATGTCACATAAATTGAGAAAGCTGACTTACTA 240
Qy 828 AAGGCTATTCAAGAACATTGATGCTTAAGCTCAAGTAAAGCACTTAAGTCAAGTTC 887
Db 241 AAGGCTATTCAAGAACATTGATGCTTAAGCTCAAGTAAAGCACTTAAGTCAAGTTC 300
Qy 888 ATTGATTGTCAGACGATGCAACCATTAATGATGCAAAAGGCAAGTCTACTTGTGAC 947
Db 301 ATTGATTGTCAGACGATGCAACCATTAATGATGCAAAAGGCAAGTCTACTTGTGAC 360
Qy 948 AAAGATGTTGCGTAACCTTGGCCGACCCACCTGTCAGAAATTTTGTAGGCGACAT 1007
Db 361 AAAGATGTTGCGTAACCTTGGCCGACCCACCTGTCAGAAATTTTGTAGGCGACAT 420
Qy 1008 GTGCGCGTTAGACCATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 1067
Db 421 GTGCGCGTTAGACCATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 480
Qy 1068 GAAATATACGTGACGTTACTCCCTTAAACCCTGATGACGATTTTCAAGCAGTCTCAA 1127
Db 481 GAAATATACGTGACGTTACTCCCTTAAACCCTGATGACGATTTTCAAGCAGTCTCAA 540
Qy 1128 GATATACGATTTTGAAGAAACATGATGCTGATGACACATCACTCAAGAAATTTACTA 1187
Db 541 GATATACGATTTTGAAGAAACATGATGCTGATGACACATCACTCAAGAAATTTACTA 600
Qy 1188 GCTCAAGCAAGAGCATTTTAAACAAAGAACCCAGGCTATAGATTTTGAAGACGTAC 1247
Db 601 GCTCAAGCAAGAGCATTTTAAACAAAGAACCCAGGCTATAGATTTTGAAGACGTAC 660
Qy 1248 TCCTCAATGCTCACTGACCAATGACATTTTCGTGAGATTTTCAAGTCAAGTCAAG 1307
Db 661 TCCTCAATGCTCACTGACCAATGACATTTTCGTGAGATTTTCAAGTCAAGTCAAG 720
Qy 1308 TTTACTTACCGGTGTTAAATTCGGGAAACAGCTTATAGATCAATAAATTCGTGTG 1367
Db 721 TTTACTTACCGGTGTTAAATTCGGGAAACAGCTTATAGATCAATAAATTCGTGTG 780
Qy 1368 AATGAGAAATTAACCAACATGACCTGATCTGAGAAATTTACGTCTTAAAAAGGG 1427
Db 781 AATGAGAAATTAACCAACATGACCTGATCTGAGAAATTTACGTCTTAAAAAGGG 840
Qy 1428 GAAAGGCGGTATGATCCCTTGAATCGAGTCACTTGAACCTGTCACCATCAATAGCT 1487
Db 841 GAAAGGCGGTATGATCCCTTGAATCGAGTCACTTGAACCTGTCACCATCAATAGCT 900
Qy 1488 GATGTCATCAAGCAAGATGCTTAAGAAAGTGAAGAGCTTAAAGCTTAAGGAAAGCT 1547
Db 901 GATGTCATCAAGCAAGATGCTTAAGAAAGTGAAGAGCTTAAAGCTTAAGGAAAGCT 960
Qy 1548 TTTGACTTCAGAGATTTATAGATCTGCTGATTAAGGCTAAAGCTTCAACAAATCTC 1607
Db 961 TTTGACTTCAGAGATTTATAGATCTGCTGATTAAGGCTAAAGCTTCAACAAATCTC 1020
Qy 1608 GATGCTTTGATTAATGAGCTAATCCTTAACTGGAAGAAAGTAAAGAGTAAACAGATGAC 1667
Db 1021 GATGCTTTGATTAATGAGCTAATCCTTAACTGGAAGAAAGTAAAGAGTAAACAGATGAC 1080
Qy 1668 ACCAACCCTATCATTAACCGTTTATATGAGGCAAGGCAAGGAAAGAGATGCTAGCTAC 1727
Db 1081 ACCAACCCTATCATTAACCGTTTATATGAGGCAAGGCAAGGAAAGAGATGCTAGCTAT 1140

Qy 1728 CATTTAGCTGTGCT 1742
Db 1141 CATTTAGCTGTATGAT 1155

RESULT 12
CQ797820 7057 bp DNA linear PAT 20-APR-2004
LOCUS CQ797820
DEFINITION Sequence 26 from Patent WO2004029256.
ACCESSION CQ797820
VERSION CQ797820.1 GI:46426093
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.

REFERENCE
1 selman-housein Sosa,G., aguiar Cabeza,E., gonz lez Quintero,A.D. and ramos gonz lez,O.
Vector for the production of transplastomic angiosperm plants
Patent: WO 2004029256-A 26 08 -APR-2004;
CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA (CU)
Location/Qualifiers
1..7057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Nucleotide sequence of DNA fragment from the vector pVPTA-Batrep between the rice atpb and tobacco rbcl borders."

FEATURES
source

ORIGIN

Query Match 53.7%; Score 1124.6; DB 6; Length 7057;
Best Local Similarity 98.4%; Pred. No. 9.9e-292;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCAATGTCATCAACAGCCAAATGATGT 647
Db 2023 ATTGCTGACCTGAGTGGCTGCTAGACCGTCAATGTCATCAACAGCCAAATGATGT 2082
Qy 648 AGCGTTGCTGATCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 707
Db 2083 AGCGTTGCTGATCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 2142
Qy 708 GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
Db 2143 GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 2202
Qy 768 AAACATTGCTACGATAGTGGCGGATGTCACATAAATTGAGAAAGCTGACTTACTA 827
Db 2203 AAACATTGCTACGATAGTGGCGGATGTCACATAAATTGAGAAAGCTGACTTACTA 2262
Qy 828 AAGGCTATTCAAGAAACAATTGATGCTTAAGCTCAAGTAAAGCACTTGTAGAGTC 887
Db 2263 AAGGCTATTCAAGAAACAATTGATGCTTAAGCTCAAGTAAAGCACTTGTAGAGTC 2322
Qy 888 ATTGATTGTCAGACGATGCAACCATTAATGATGCAAAAGGCAAGTCTACTTGTGAC 947
Db 2323 ATTGATTGTCAGACGATGCAACCATTAATGATGCAAAAGGCAAGTCTACTTGTGAC 2382
Qy 948 AAAGATGTTGCGTAACTTGGCCGACCAACCTGTCAGAAATTTTGTGAAGCGACAT 1007
Db 2383 AAAGATGTTGCGTAACTTGGCCGACCAACCTGTCAGAAATTTTGTGAAGCGACAT 2442
Qy 1008 GTGCGCGTTAGACCATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 1067
Db 2443 GTGCGCGTTAGACCATTAAGAAAGAAACCAATACAAAGCAAGGAAATCTGTGATGTG 2502
Qy 1068 GAATATACCTGATAGTTACTCCCTTAAACCTGATGACGATTTGACAGCAGTCTCAAA 1127
Db 2503 GAATATACCTGATAGTTACTCCCTTAAACCTGATGACGATTTGACAGCAGTCTCAAA 2562
Qy 1128 GATATCAAGCTATTGAAAAACATGACTATCGGTGACACATCACTCAAGAAATTTACTA 1187

Db 2563 GATACGAGGATTTGAAAAACATGATCGTATCGGTGACACCATCATCTCAAGAAATTACTA 2622
Qy 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCCAGCGCTATACGATTTATGAAACGTGAC 1247
Db 2623 GCTCAAGCAAAAGCATTTTAAACAAAACCCAGCGCTATACGATTTATGAAACGTGAC 2682
Qy 1248 TCCTCATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAG 1307
Db 2683 TCCTCATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAG 2742
Qy 1308 TTTACTTACCGTGTAAATTCGGGAAACAAGCTTATAGATCAATAAATATCGTCTG 1367
Db 2743 TTTACTTACCATGTCAAAAATTCGGGAAACAAGCTTATAGATCAATAAATATCGTCTG 2802
Qy 1368 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGG 1427
Db 2803 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAGG 2862
Qy 1428 GAAAAGCGGTATGATCCCTTGATCGAGTCACTTGAAAGCTTACCATCAATATAGTT 1487
Db 2863 GAAAAGCGGTATGATCCCTTGATCGAGTCACTTGAAAGCTTACCATCAATATAGTT 2922
Qy 1488 GATGCGATACCAACGATTTGTTAAAGTGAAGCACTTTAAAGCTAGGGAACGTAAAC 1547
Db 2923 GATGCGATACCAACGATTTGTTAAAGTGAAGCACTTTAAAGCTAGGGAACGTAAAC 2982
Qy 1548 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAACTTCAACATCTC 1607
Db 2983 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAACTTCAACATCTC 3042
Qy 1608 GATGCTTTTGTATTTATGACCTTAACTTGAAAGTGAAGTATTAACAGATGAC 1667
Db 3043 GATGCTTTTGTATTTATGACCTTAACTTGAAAGTGAAGTATTAACAGATGAC 3102
Qy 1668 ACCAACCGTATCAATACCGTTTATATGAGCAAGCAACCGGAAGAGAAATGCTAGTAC 1727
Db 3103 ACCAACCGTATCAATACCGTTTATATGAGCAAGCAACCGGAAGAGAAATGCTAGTAC 3162
Qy 1728 CATTAGCTGTGGGT 1742
Db 3163 CATTAGCTGTATGAT 3177

RESULT 13
A20006 1335 bp DNA linear PAT 14-JUN-1995
LOCUS A20006
DEFINITION SEO ID NO: 14; Streptokinase gene.
ACCESSION A20006
VERSION A20006.1 GI:1247579
KEYWORDS
SOURCE
ORGANISM
Streptococcus dysgalactiae subsp. equisimilis
Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 1335)

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.1335
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
7.1335
/codon_start=1
/transl_table=11
/product="Streptokinase"
/protein_id="CAA01484.1"
/db_xref="GI:1247580"
/translation="MKNYLSFGMPALIFALTPTGVNVSVOALAGPEWLLDPSVNSQL
VSVAGTIVSGTNDISLKPEIDLTSPAHGKTBOGLSPKSPFATDSGAMPKLER

CDS

ORIGIN

Query Match 53.5%; Score 1122.4; DB 6; Length 1335;
Best Local Similarity 98.2%; Pred. No. 3.4e-291;
Matches 1135; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ADULKALIOELIANVHNSNDYFEVIEDASDATITDRNGKVFADKDSVTLPTQPVOR
FLISGVRVRYKEKPIONOKASVDVETVQFPLANDPDDPRPLKDTKLKTLAIGD
TINSOELAAOSILNTHPGVTVYEDSSIVTHDNIDIPRILFPMDEFTYVNRBO
AYEINKSGLANEINRNDILSEKTYVKKEKSPDPDRSHLKLFTIKYDVNRNELL
KSEQLTASERNLFRDLYDRDAKLLYNLDAFGLMDYTLTKVEVDNRNLLIT
VYMKRBEGENASYHALYDORYTEEREVYSYLRVGTPIPDNPDK"

Qy 587 TATTGTGACCTGAGTGGCTGTAGACCGGCACTGTCAACAACAGCCAAATGGTGT 646
Db 84 TATTGTGACCTGAGTGGCTGTAGACCGGCACTGTCAACAACAGCCAAATGGTGT 143
Qy 647 TAGCGTGTGGTATCTGTTGAGGGGACGAATCAAGAATTAGCTTAAATTTTGAAT 706
Db 144 TAGCGTGTGGTATCTGTTGAGGGGACGAATCAAGAATTAGCTTAAATTTTGAAT 203
Qy 707 CGATCTAACATCAAGACTGCTCATGAGAGAAAGACAGAGAGCTTAAGTCCAAATC 766
Db 204 TGACCTAACATCAAGACTGCTCATGAGAGAAAGACAGAGAGCTTAAGTCCAAATC 263
Qy 767 AAAACCAATTCCTAGTAGTAGGCGGAGTGCATTAACCTTGAGAAAGCTACTACT 826
Db 264 AAAACCAATTCCTAGTAGTAGGCGGAGTGCATTAACCTTGAGAAAGCTACTACT 323
Qy 827 AAAAGCTATTCAGAAACAATTGATCGCTAACGTCACAGTAACGACCTTATGAGGT 886
Db 324 AAAAGCTATTCAGAAACAATTGATCGCTAACGTCACAGTAACGACCTTATGAGGT 383
Qy 887 CATTGATTTTGCAAGCGATGCAACATTAATCTGATGCAAGCGCAAGGTCTACTTGTCTGA 946
Db 384 CATTGATTTTGCAAGCGATGCAACATTAATCTGATGCAAGCGCAAGGTCTACTTGTCTGA 443
Qy 947 CAAGATGTTGGTGGTAACTTGGCGGACCGCAACCTGTCCAGAAATTTTGTCAAGCGCA 1006
Db 444 CAAGATGTTGGTGGTAACTTGGCGGACCGCAACCTGTCCAGAAATTTTGTCAAGCGCA 503
Qy 1007 TGTGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGCGAATCTGTGATGT 1066
Db 504 TGTGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGCGAATCTGTGATGT 563
Qy 1067 GGAATATATCTGTACAGTTTACTCCCTTAACTCTGATGACGATTTCAAGCAAGTCTCAA 1126
Db 564 GGAATATATCTGTACAGTTTACTCCCTTAACTCTGATGACGATTTCAAGCAAGTCTCAA 623
Qy 1127 AGATATCTAAGCTATTTGAAAAACATAGCTATGCGTGAACAACATCACTCAAGAATTACT 1186
Db 624 AGATATCTAAGCTATTTGAAAAACATAGCTATGCGTGAACAACATCACTCAAGAATTACT 683
Qy 1187 AGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATGCAATTTTGAACGTGA 1246
Db 684 AGCTCAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATGCAATTTTGAACGTGA 743
Qy 1247 CTCCTCAATGCTCATGACGAATGACATTTTCCGTACGATTTTCAATGATGATCAAGA 1306
Db 744 CTCCTCAATGCTCATGACGAATGACATTTTCCGTACGATTTTCAATGATGATCAAGA 803
Qy 1307 GTTTACTTACCGTGTATTAATAATTCGGGAACAAGCTTATAGATCAATAAATATCGGTCT 1366
Db 804 GTTTACTTACCGTGTATTAATAATTCGGGAACAAGCTTATAGATCAATAAATATCGGTCT 863
Qy 1367 GATGAAGAATAAACAACATGACCTGATCTGAGAAATATTAAGTCTTTAAAAAAGG 1426
Db 864 GATGAAGAATAAACAACATGACCTGATCTGAGAAATATTAAGTCTTTAAAAAAGG 923
Qy 1427 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACGTGTCAACATCAATAGCT 1486
Db 924 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACGTGTCAACATCAATAGCT 983
Qy 1487 TGATGTGATTCACCAAGCAATTCCTTAAAAAGTGAAGCACTCTTAAACAGCTAGCAACGTAA 1546


```

/codon_start=1
/translation=1
/product="streptokinase fused to yeast alpha-factor"
/db_xref="GI:1247851"
/translation="MRPPIFTAVLPAASSALAPVNTTEDETAQIPAEAVIGYLDL
EGPFDVAVLPFSSTNNGLFINTTASIAKEBGSVLDKRIAPEMILDPSPVNSQ
LVSVAGTEGNTQDILKFEIDLTSPAPGKTEQGLSPKSPAPATDSGAMPKLE
KADLKAIOROLIANVNSNDYEPVIDPASPATTDNGKYPADKXSVLPPOPC
EPILSGHVRVRYKPKRIQONAKSVDPYQPTPLPDDDFRGLKDTKLKTLAIG
DITTSOELIAQOSILNKTHPEGTYIYERDSSIVHNDIFPTLIPMOEFTYHAKNE
QAYEINKSGSLNEINNTDILSEKXYVAKKEKPYDDEPDRSHLFLPIKYDVNTNEL
LKESOLLTASERINDFRDLDPDRKAKILYNNLDAPFGIMDYTLGKXEDNDDTNRIL
TYVMGKRPBENASVHLAYDKRYTEEREERYSYLRYTGPTIPNPMDK"

```

ORIGIN

```

Query Match      53.5%; Score 1122.4; DB 6; Length 1512;
Best Local Similarity 97.8%; Pred. No. 3.4e-291;
Matches 1138; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 579 GATGTTGATATTGCTGAGACCTGAGTGGCTGTAGACCTCATCTGTCAACAACAGCCAA 638
DB 253 GATTAAGAAATGCTGAGACCTGAGTGGCTGTAGACCTCATCTGTCAACAACAGCCAA 312
QY 639 TTGGTTGTAGCGTTGCTGCTGCTGTTAGAGGGAAGCAATCAAGACATTAGCTTAAATTT 698
DB 313 TTAGTTCTTAGCGTTGCTGCTGCTGTTAGAGGGAAGCAATCAAGACATTAGCTTAAATTT 372
QY 699 TTTGAAATCGATCTAATCAACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAAT 758
DB 373 TTTGAAATTCGATCAATCAACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAAT 432
QY 759 CCAAAATCAAAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
DB 433 CCAAAATCAAAACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 492
QY 819 GACTTACTAAAGGCTATTCAAGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878
DB 493 GACTTACTAAAGGCTATTCAAGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
QY 879 TTTGAGGTCAATTGATTTTGGCAAGGATGCAACCAATTAATGATCAACCAAGGCTTAC 938
DB 553 TTTGAGGTCAATTGATTTTGGCAAGGATGCAACCAATTAATGATCAACCAAGGCTTAC 612
QY 939 TTTGCTGCAAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
DB 613 TTTGCTGCAAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
QY 999 AGCGGACATGAGCGCTTGAACCATATTAAGAAACCAATCAAAACCAAGCGAAATCT 1058
DB 673 AGCGGACATGAGCGCTTGAACCATATTAAGAAACCAATCAAAACCAAGCGAAATCT 732
QY 1059 GTTGATGTGAATTAATCTGTACAGTTTACTCCCTTAAACCTGTATGAGATTTGAGCA 1118
DB 733 GTTGATGTGAATTAATCTGTACAGTTTACTCCCTTAAACCTGTATGAGATTTGAGCA 792
QY 1119 GGTCTCAAAGATTAAGTATTGAAAACTAGCTATCGGTGACACCATCAATCTCA 1178
DB 793 GGTCTCAAAGATTAAGTATTGAAAACTAGCTATCGGTGACACCATCAATCTCA 852
QY 1179 GAATTAAGTCTAAGCAAAAGATTTTAAACAAACCAAGGCTATAGCAATTTAT 1238
DB 853 GAATTAAGTCTAAGCAAAAGATTTTAAACAAACCAAGGCTATAGCAATTTAT 912
QY 1239 GAAAGTGAATCTCAATCTGCTCATGATGACATTTTCCGTACGATTTTACCAATG 1298
DB 913 GAAAGTGAATCTCAATCTGCTCATGATGACATTTTCCGTACGATTTTACCAATG 972
QY 1299 GATCAAGAGTTTACTTACCGGTGTTAAATCGGAAACAAAGCTTATAGGATCAATTA 1358
DB 973 GATCAAGAGTTTACTTACCGGTGTTAAATCGGAAACAAAGCTTATAGGATCAATTA 1032
QY 1359 TCTGCTGATGAAGAAATTAACAACGATGATCTGTGAGAAATATTAGTCTT 1418

```

```

DB 1033 TCTGCTGATGAAGAAATTAACAACGATGATCTGTGAGAAATATTAGTCTT 1092
QY 1419 AAAAAAGGGAAGAAAGCGGTATGATCCCTTTATGTGACATCTTGAACCTGTACCAATC 1478
DB 1093 AAAAAAGGGAAGAAAGCGGTATGATCCCTTTATGTGACATCTTGAACCTGTACCAATC 1152
QY 1479 AATATCGTGAATGCTGATTAACAAGAAATGCTTAAAGAGTGAAGCTCTTAACAGCTAGC 1538
DB 1153 AATATCGTGAATGCTGATTAACAAGAAATGCTTAAAGAGTGAAGCTCTTAACAGCTAGC 1212
QY 1539 GAAAGTGAATCTTGAATCTGAGATTTATACGATCCCTGATTAAGGCTTAACTCTTAC 1598
DB 1213 GAAAGTGAATCTTGAATCTGAGATTTATACGATCCCTGATTAAGGCTTAACTCTTAC 1272
QY 1599 AACAATCTGATGCTTTTGTATTAATGATTAATCTTAATCTGAAAAATAGAGATAT 1658
DB 1273 AACAATCTGATGCTTTTGTATTAATGATTAATCTTAATCTGAAAAATAGAGATAT 1332
QY 1659 CACGATGACACCAACCGATATGATTAATGATTAATGATTAATGATTAATGATTAAT 1718
DB 1333 CACGATGACACCAACCGATATTAATGATTAATGATTAATGATTAATGATTAATGAT 1392
QY 1719 GCTAGCTACCAATTTAGCTGTGCT 1742
DB 1393 GCTAGCTACCAATTTAGCTGTGCT 1416

```

Search completed: February 1, 2006, 04:17:49
 Job time : 10656.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 10:44:08 ; Search time 1267.07 Seconds
(without alignment)
11024.772 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 2096 1 cgaagaccatcatgtgtc.....cttcaccgatgtcgttag 2096

Sequence: 1 cgaagaccatcatgtgtc.....cttcaccgatgtcgttag 2096

Scoring table: IDENTITY NUC
Gapop 10'-0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues 9993994

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2096	100.0	2096	3	AAA37643
2	1684	80.3	1782	3	AAA37642
3	1496.2	71.4	1541	3	AAA37644
4	1165.8	55.6	1661	3	AAA37637
5	1150.8	54.9	2385	2	AAx80497
6	1147	54.7	1377	3	AAA37622
7	1145.4	54.6	1242	2	AAx80492
8	1145.4	54.6	1245	3	AAA37633
9	1145.4	54.6	1254	6	ABA05546
10	1145.4	54.6	8893	6	ABA05547
11	1143.8	54.6	1242	2	AAx16632
12	1143.8	54.6	2030	2	AAQ11651
13	1140.6	54.4	1242	2	AAx16633
14	1134.2	54.1	1327	3	AAA37628
15	1132.4	54.0	2366	2	AAx77778
16	1125.4	53.7	1458	2	AAQ12162
17	1124.6	53.7	7057	12	ADW01294
18	1122.4	53.5	1335	2	AAQ12156
19	1122.4	53.5	1512	2	AAQ12158

20	1122.2	53.5	2589	2	AAQ12160	AaQ12160 OmpA str
21	1121.4	53.5	1245	2	AAQ20665	AaQ20665 SKC-2 str
22	1121.4	53.5	1467	2	AAQ12490	AaQ12490 Factor Xa
23	1116.6	53.3	1407	1	AAx70106	Aa70106 DNA encod
24	1114.4	53.2	1473	2	AAQ05603	AaQ05603 Sequence
25	1106.4	52.8	2568	1	AAx50493	Aa50493 Vector ps
26	1104.8	52.7	1323	2	AAx29961	Aa29961 Mutant st
27	1096.2	52.3	1442	5	AAx82144	Aa82144 DNA encod
28	1093.6	52.2	1122	3	AAx299251	Aa299251 DNA encod
29	1093.6	52.2	1158	3	AAx299252	Aa299252 DNA encod
30	1086	51.8	1209	3	AAx299249	Aa299249 DNA encod
31	1086	51.8	1245	3	AAx299250	Aa299250 DNA encod
32	1077.2	51.4	2253	2	AAQ12161	AaQ12161 Met-core
33	1076	51.3	1119	2	AAQ12159	AaQ12159 Truncated
34	996	47.5	1320	6	ABx70192	ABx70192 Streptoco
35	996	47.5	1323	13	ADx83811	ADx83811 S. pyogen
36	991.2	47.3	1473	2	AAQ05604	AaQ05604 Streptoki
37	973	46.4	1068	2	AAx80493	Aa80493 Recombina
38	968.4	46.2	2208	2	AAx83589	Aa83589 Recombina
39	899.4	42.9	1245	10	ADx48644	ADx48644 Streptoco
40	759.4	36.2	1262	2	AAQ10230	AaQ10230 Synthetic
41	450	21.5	450	2	AAx80494	Aa80494 Recombina
42	330	15.7	777	3	AAx37632	Aa37632 Human fib
43	330	15.7	1929	13	ADx17488	ADx17488 Nucleotid
44	330	15.7	1929	13	ADx97657	ADx97657 Human fib
45	330	15.7	1929	14	ADx44479	ADx44479 Human fib

ALIGNMENTS

RESULT 1	AAA37643	standard; DNA; 2096 BP.
ID	AAA37643	
XX	AAA37643;	
AC	15-SEP-2003 (revised)	
XX	13-OCT-2000 (first entry)	
DT	Chimeric SK-FBD coding sequence.	
XX	Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;	
KW	plasminogen; human; fibrinectin; thrombolytic therapy;	
KW	cardiovascular disorder; ss.	
XX	Streptococcus dysgalactiae subsp. equisimilis.	
OS	Homo sapiens.	
OS	Chimeric.	
XX	EP1024192-A2.	
XX	02-AUG-2000.	
XX	23-DEC-1999; 99BP-00310541.	
PF	24-DEC-1998; 98IN-DE003825.	
XX	(COUL) CSIR COUNCIL SCI IND RES.	
PA	Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;	
PI	Yadav M;	
PI	WPI, 2000-516032/47.	
XX	Hybrid streptokinase-fibrin binding domain polypeptides useful for	
PT	thrombolytic therapy comprises a streptokinase fused with fibrin binding	
PT	domains of human fibronectin.	
XX	Example 6; Fig 22b; 58pp; English.	
PS	This sequence represents a chimeric streptokinase-fibrin binding domain	
XX	(SK-FBD) protein coding sequence. The invention relates to a hybrid	
CC		

CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinogen, which are capable of
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-Sep-2003 to standardise OS field)

XX Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;

Query Match 100.0%; Score 2096; DB 3; Length 2096;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAAGACCATTCATGTTGTTGCTCGAGTCGACGTTTTCAGACAGCTCGCTTCAGC 60
DB 1 CGAAGACCATTCATGTTGTTGCTCGAGTCGACGTTTTCAGACAGCTCGCTTCAGC 60
QY 61 TTGCGTCGCGTATCGGATGTTCACTTCGCTAACAGTAAGGCAACCCCGCAGCTTAGCC 120
DB 61 TTGCGTCGCGTATCGGATGTTCACTTCGCTAACAGTAAGGCAACCCCGCAGCTTAGCC 120
QY 121 GGGTCTTCAACGACAGGACGATCATGCGCACCCGTGGCAGAGCCCAACGCTGCGCG 180
DB 121 GGGTCTTCAACGACAGGACGATCATGCGCACCCGTGGCAGAGCCCAACGCTGCGCG 180
QY 181 AGATCTCGATCCGCGAATTAATACGATCATATAGGAGAACACACGCTTCCCTC 240
DB 181 AGATCTCGATCCGCGAATTAATACGATCATATAGGAGAACACACGCTTCCCTC 240
QY 241 TAGAATTAATTTTGTTAACCTTAAGAGAGATATACCATGCTGCAAGCAACAGTT 300
DB 241 TAGAATTAATTTTGTTAACCTTAAGAGAGATATACCATGCTGCAAGCAACAGTT 300
QY 301 GTACCATAGCTGAGAAAGTGTTCATCATGCTGCTGGAACTTCCTATGCTGCGAGAA 360
DB 301 GTACCATAGCTGAGAAAGTGTTCATCATGCTGCTGGAACTTCCTATGCTGCGAGAA 360
QY 361 ACCTGGAGAGAGGACGCGAGCATCACTTCTTAAGAAATAGATGACAGATCAGG 420
DB 361 ACCTGGAGAGAGGACGCGAGCATCACTTCTTAAGAAATAGATGACAGATCAGG 420
QY 421 ACACAAGACATCTTATAGATTGAGACACCTGAGCAAGAGAGATTAATGAGAGAAC 480
DB 421 ACACAAGACATCTTATAGATTGAGACACCTGAGCAAGAGAGATTAATGAGAGAAC 480
QY 481 TGCTCAGAGCATCTGACAGGCAAGGCGAGAGAGTGAAGTGAAGAGGACACCT 540
DB 481 TGCTCAGAGCATCTGACAGGCAAGGCGAGAGAGTGAAGTGAAGAGGACACCT 540
QY 541 CTGTGAGACACATCAGAGGAGCTGCGCCCTTCAACGATGTTGATTTGCTGAGCCTG 600
DB 541 CTGTGAGACACATCAGAGGAGCTGCGCCCTTCAACGATGTTGATTTGCTGAGCCTG 600
QY 601 AGTGGCTGCTAGACCGTTCATCTGTCAACACAGCCAAATGTTGTTAGGCTGCTGTA 660
DB 601 AGTGGCTGCTAGACCGTTCATCTGTCAACACAGCCAAATGTTGTTAGGCTGCTGTA 660
QY 661 CTGTGAGGAGAGATCAAGACATTAAGTCTTAATTTTGAATGATCTTAACATCAC 720
DB 661 CTGTGAGGAGAGATCAAGACATTAAGTCTTAATTTTGAATGATCTTAACATCAC 720

DB 661 CTGTGAGGAGAGATCAAGACATTAAGTCTTAATTTTGAATGATCTTAACATCAC 720
QY 721 GACCTGCTCAGAGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAACATTTGCTA 780
DB 721 GACCTGCTCAGAGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAACATTTGCTA 780
QY 781 CTGATAGGCGGCGATGTCATATAAATTGAGAAAGCTGACTTAAGGCTATTCAAG 840
DB 781 CTGATAGGCGGCGATGTCATATAAATTGAGAAAGCTGACTTAAGGCTATTCAAG 840
QY 841 AACATTAATGCTTAACGTCACAGTAACGACGTAACGTAACGTAACGTAACGTAAC 900
DB 841 AACATTAATGCTTAACGTCACAGTAACGACGTAACGTAACGTAACGTAACGTAAC 900
QY 901 GCGATGCAACCATTAATGATGAAACGCGAAGGTTACTTTGCTGACAAAGTGGTGG 960
DB 901 GCGATGCAACCATTAATGATGAAACGCGAAGGTTACTTTGCTGACAAAGTGGTGG 960
QY 961 TTAACCTGCGAGCCCAACCTGTCACAGAAATTTTGTGAAGGCGACATGCGCGTTAGC 1020
DB 961 TTAACCTGCGAGCCCAACCTGTCACAGAAATTTTGTGAAGGCGACATGCGCGTTAGC 1020
QY 1021 CATATTAAGAAAGAAACCATATACAAACCAAGGAAATCTGTGATGTAATTAATCTGAC 1080
DB 1021 CATATTAAGAAAGAAACCATATACAAACCAAGGAAATCTGTGATGTAATTAATCTGAC 1080
QY 1081 AGTTTACTCCTTAAACCTGATGACGATTTCAAGCCAGGCTCAAAATTAAGTAT 1140
DB 1081 AGTTTACTCCTTAAACCTGATGACGATTTCAAGCCAGGCTCAAAATTAAGTAT 1140
QY 1141 TGAAGAACATAGCTATGCGTACACCATCAATCTCAAGATTTACTCAAGACAA 1200
DB 1141 TGAAGAACATAGCTATGCGTACACCATCAATCTCAAGATTTACTCAAGACAA 1200
QY 1201 GCATTTTAAACAAAGAAACCAACGAGCTTAACGATTTATGAACGTAACCTCAATCTCA 1260
DB 1201 GCATTTTAAACAAAGAAACCAACGAGCTTAACGATTTATGAACGTAACCTCAATCTCA 1260
QY 1261 CTGATGACATGACATTTTCCGTAGATTTTACCAATGATCAAGAGTTTCTTACCGTG 1320
DB 1261 CTGATGACATGACATTTTCCGTAGATTTTACCAATGATCAAGAGTTTCTTACCGTG 1320
QY 1321 TTTAAATTCGGGAAACAGCTTAATAGATCAATTAATAATCTGCTGAATGAAGAAATTA 1380
DB 1321 TTTAAATTCGGGAAACAGCTTAATAGATCAATTAATAATCTGCTGAATGAAGAAATTA 1380
QY 1381 ACAACACGACCTGATCTCTGAGAAATTTAGCTCTTAAAGAGGAGAGGAGGAGGATG 1440
DB 1381 ACAACACGACCTGATCTCTGAGAAATTTAGCTCTTAAAGAGGAGAGGAGGAGGATG 1440
QY 1441 ATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACATCAAAATACGTTGATGATACCA 1500
DB 1441 ATCCCTTTGATGCGAGTCACTTGAACCTGTTCAACATCAAAATACGTTGATGATACCA 1500
QY 1501 ACGAATGCTTAAGAGAGAGCTCTTAACAGCTAGAGAAAGCTTAACCTTAACAGAG 1560
DB 1501 ACGAATGCTTAAGAGAGAGCTCTTAACAGCTAGAGAAAGCTTAACCTTAACAGAG 1560
QY 1561 ATTTATAGATCTCTGATGTAAGAGCTTAAGCTTAACCAACATCTGATGCTTTGCTA 1620
DB 1561 ATTTATAGATCTCTGATGTAAGAGCTTAAGCTTAACCAACATCTGATGCTTTGCTA 1620
QY 1621 TTATGACTATATCTTAACGTAAGAAAGTGAAGATTAACAGATGACACCAACGCTATCA 1680
DB 1621 TTATGACTATATCTTAACGTAAGAAAGTGAAGATTAACAGATGACACCAACGCTATCA 1680
QY 1681 TTAACCTTTATATGAGCAAGGACCCGAGAGAGAGATGCTTAACCTTAACCTTAAGT 1740
DB 1681 TTAACCTTTATATGAGCAAGGACCCGAGAGAGAGATGCTTAACCTTAACCTTAAGT 1740
QY 1741 GTGCGCAGGCGCAACAGTTGTAACCAATAGCTGAAGATGTTTGAATCATGCTGCTGGA 1800
DB 1741 GTGCGCAGGCGCAACAGTTGTAACCAATAGCTGAAGATGTTTGAATCATGCTGCTGGA 1800

QY 1801 CTTCTATGTTGTCGAGAAACGTGGAGAAACCCCTACCAAGCTGTGATGTAGATT 1860
CC CTTCTATGTTGTCGAGAAACGTGGAGAAACCCCTACCAAGCTGTGATGTAGATT 1860
Db 1801 CTTCTATGTTGTCGAGAAACGTGGAGAAACCCCTACCAAGCTGTGATGTAGATT 1860
QY 1861 GTACTTGTCCCTGGAGAAAGCGACGCGATCATCTTGACATTTTGAAATAGATCAACG 1920
CC GTACTTGTCCCTGGAGAAAGCGACGCGATCATCTTGACATTTTGAAATAGATCAACG 1920
Db 1861 GTACTTGTCCCTGGAGAAAGCGACGCGATCATCTTGACATTTTGAAATAGATCAACG 1920
QY 1921 ATCAGACACAGAAGACATCTCTATAGAAATTGAGACACCTGGAGCAAGAAAGTAAATCGAG 1980
CC ATCAGACACAGAAGACATCTCTATAGAAATTGAGACACCTGGAGCAAGAAAGTAAATCGAG 1980
Db 1921 ATCAGACACAGAAGACATCTCTATAGAAATTGAGACACCTGGAGCAAGAAAGTAAATCGAG 1980
QY 1981 GAAACCTGCTCCAGTGATCTGCAACAGGCACCGCCGAGAGAGATGTAAGTGTAGAGGC 2040
CC GAAACCTGCTCCAGTGATCTGCAACAGGCACCGCCGAGAGAGATGTAAGTGTAGAGGC 2040
Db 1981 GAAACCTGCTCCAGTGATCTGCAACAGGCACCGCCGAGAGAGATGTAAGTGTAGAGGC 2040
QY 2041 ACACCTCTGTGACAGACCATCGACGCGATCTGGCCCTTCACCGATGTCTGTTAG 2096
CC ACACCTCTGTGACAGACCATCGACGCGATCTGGCCCTTCACCGATGTCTGTTAG 2096
Db 2041 ACACCTCTGTGACAGACCATCGACGCGATCTGGCCCTTCACCGATGTCTGTTAG 2096

RESULT 2
AAA37642 standard; DNA; 1782 BP.
ID AAA37642 standard; DNA; 1782 BP.
XX
AC AAA37642;
XX
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
DE
DE Chimeric SK-FBD coding sequence.
XX
KM Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KM plasminogen; human; fibrinectin; thrombolytic therapy;
KM cardiovascular disorder; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
OS
PN BP1024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
XX
PR 24-DEC-1998; 98IN-DE003825.
XX
PA (COUL) CSIR COUNCIL SCI IND RES.
XX
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
DR WPI; 2000-516032/47.
XX
PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibronectin.
XX
PS Example 5; Fig 21b; 58pp; English.
XX
CC This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibronectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of

CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolytics). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
CC
XX
SQ Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;
Query Match 80.3%; Score 1684; DB 3; Length 1782;
Beet Local Similarity 99.7%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 51 TCGCTTACAGTTTCGTCGCGTATCCGATATTCATTTCTGTAACAGTAAGCAACCCGC 110
Db 1 TCGCTTACAGTTTCGTCGCGTATCCGATATTCATTTCTGTAACAGTAAGCAACCCGC 60
QY 111 CAGCTTACCGGGTCTCAACGACAGAGACGATCATGCGACCCGTGGCCAGACCCA 170
Db 61 CAGCTTACCGGGTCTCAACGACAGAGACGATCATGCGACCCGTGGCCAGACCCA 120
QY 171 ACGTCCCGGAGATTCGATCCCGGAATTAATGACATCACTATAGGAGAACCAAC 230
Db 121 ACGTCCCGGAGATTCGATCCCGGAATTAATGACATCACTATAGGAGAACCAAC 180
QY 231 GATTTCCTCTAGAAATTAATTTTGTTAACCTTAAGAGAGATTAACAGATGTCACAGC 290
Db 181 GATTTCCTCTAGAAATTAATTTTGTTAACCTTAAGAGAGATTAACAGATGTCACAGC 240
QY 291 ACAACAGATTAACCATAGCTGAGAGAGTGTATGATCATGCTGTGGACTTCTATGT 350
Db 241 ACAACAGATTAACCATAGCTGAGAGAGTGTATGATCATGCTGTGGACTTCTATGT 300
QY 351 GGTGGAGAAACGTGGAGAAAGGACGCGAGCATCATCTTGACATTTTGAAATGATGC 410
Db 301 GGTGGAGAAACGTGGAGAAAGGACGCGAGCATCATCTTGACATTTTGAAATGATGC 360
QY 411 AAGCATCAGACACAGAGACATCCATATGAAATTGAGACACCTGAGCAAGAAAGATAT 470
Db 361 AAGCATCAGACACAGAGACATCCATATGAAATTGAGACACCTGAGCAAGAAAGATAT 420
QY 471 CGAGGAAACCTGCTCCAGTGCATCTGCAACGACGACGCGAGAGAGTGAAGTGTAG 530
Db 421 CGAGGAAACCTGCTCCAGTGCATCTGCAACGACGACGCGAGAGAGTGAAGTGTAG 480
QY 531 AGGCACACCTCTGTGACAGACCATCGAGCGGATCTGGCCCTTCACCGATGTTGTAAT 590
Db 481 AGGCACACCTCTGTGACAGACCATCGAGCGGATCTGGCCCTTCACCGATGTTGTAAT 540
QY 591 GCTGACCTGAGTGGCTGTGACCGTTCATCTGTGCAACACAGCCATTTGTTTATGC 650
Db 541 GCTGACCTGAGTGGCTGTGACCGTTCATCTGTGCAACACAGCCATTTGTTTATGC 600
QY 651 GTTGCTGTAATGTTGAGGGAGCAAGATCAAGACATTAATTTTGAATTCAT 710
Db 601 GTTGCTGTAATGTTGAGGGAGCAAGATCAAGACATTAATTTTGAATTCAT 660
QY 711 CTAACATCAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCAAAATCAAA 770
Db 661 CTAACATCAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCAAAATCAAA 720
QY 771 CCAATTTGCTACTGATTAAGTGGCGCATGTGACATTAATCTTGAGAAAGCTGACTTAAG 830
Db 721 CCAATTTGCTACTGATTAAGTGGCGCATGTGACATTAATCTTGAGAAAGCTGACTTAAG 780
QY 831 GCTATTAAGAAACAATTGATGCTTAAGTCAAGTACGACGATCACTTTGAAGTCAAT 890
Db 781 GCTATTAAGAAACAATTGATGCTTAAGTCAAGTACGACGATCACTTTGAAGTCAAT 840

```

QY      891 GATTTGCAAGCATGCAACCATTAAGTATGAGAAACGGCAAGGTCTACTTTGCTGACAA 950
      |||
DB      841 GATTTGCAAGCATGCAACCATTAAGTATGAGAAACGGCAAGGTCTACTTTGCTGACAA 900
QY      951 GATGTTGGTAACTTTGCGGACCCAACTGTCCAGAAATTTTGTAAAGCGGACATGTG 1010
      |||
DB      901 GATGTTGGTAACTTTGCGGACCCAACTGTCCAGAAATTTTGTAAAGCGGACATGTG 960
QY      1011 CGGTTAGACATTAATAAGAAAAACAATCAAAAACCAAGCAATCTGTTGATGTGAA 1070
      |||
DB      961 CGGTTAGACATTAATAAGAAAAACAATCAAAAACCAAGCAATCTGTTGATGTGAA 1020
QY      1071 TATAGTATACGTTTACTCCCTTAAACCTGATGACGATTTTCAACACAGGCTCAAAAGAT 1130
      |||
DB      1021 TATAGTATACGTTTACTCCCTTAAACCTGATGACGATTTTCAACACAGGCTCAAAAGAT 1080
QY      1131 ACTAGCTATTTGAAAACTAGCTATGATGACACATCAATCTCAAGATTAATACTAGCT 1190
      |||
DB      1081 ACTAGCTATTTGAAAACTAGCTATGATGACACATCAATCTCAAGATTAATACTAGCT 1140
QY      1191 CAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAAGTGTCTCC 1250
      |||
DB      1141 CAAGCAAAAGCATTTTAAACAAAACCAACCCAGGCTATACGATTTATGAAGTGTCTCC 1200
QY      1251 TCATGTCATCTGATGACATGACATTTTCCGTAACAATTTTAAACCAAGGATCAAGAGTTT 1310
      |||
DB      1201 TCATGTCATCTGATGACATGACATTTTCCGTAACAATTTTAAACCAAGGATCAAGAGTTT 1260
QY      1311 ACTTACCGTGTAAATCGGGAAACAAGTTATAGATCAATAAATACTGCTGAT 1370
      |||
DB      1261 ACTTACCGTGTAAATCGGGAAACAAGTTATAGATCAATAAATACTGCTGAT 1320
QY      1371 GAAGAATAAACAACACTGACCTGATCTGGAATAATTAAGTCTTAAAAAGGGAA 1430
      |||
DB      1321 GAAGAATAAACAACACTGACCTGATCTGGAATAATTAAGTCTTAAAAAGGGAA 1380
QY      1431 AAGCGGTATGATCCCTTGAATCGGACATCTGAAACTGTTCACATCAATATAGTTGAT 1490
      |||
DB      1381 AAGCGGTATGATCCCTTGAATCGGACATCTGAAACTGTTCACATCAATATAGTTGAT 1440
QY      1491 GTGATATCAACAAGATTTGTAATAAGAGAGAGCTTTAAACAGCTAGCGAACTTAACCTTA 1550
      |||
DB      1441 GTGATATCAACAAGATTTGTAATAAGAGAGAGCTTTAAACAGCTAGCGAACTTAACCTTA 1500
QY      1551 GACTTCAGAGATTTATACGATCTCGTATAGGCTTAACTACTCTCAACAATCTCGAT 1610
      |||
DB      1501 GACTTCAGAGATTTATACGATCTCGTATAGGCTTAACTACTCTCAACAATCTCGAT 1560
QY      1611 GCTTTGGTATTTATGACATTAACCTTAACTGAAAAAGTAAAGATTAACAGATGACACC 1670
      |||
DB      1561 GCTTTGGTATTTATGACATTAACCTTAACTGAAAAAGTAAAGTAAATCAAGATGACACC 1620
QY      1671 AACCGTATCAAAACCGTTTATATAGGAGAGAACCGAAGAGAGAAATGCTAGCTACAT 1730
      |||
DB      1621 AACCGTATCAAAACCGTTTATATAGGAGAGAACCGAAGAGAGAAATGCTAGCTACAT 1680
QY      1731 TTAGCTGTGTGT 1742
      |||
DB      1681 TTAGCTGTGTGT 1692

```

RESULT 3
 AAA37644
 ID AAA37644 standard; DNA; 1541 BP.
 XX

AC AAA37644;

XX 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE Chimeric SK-FBD coding sequence.
 XX

```

KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KM plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
PN EPI024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
XX
PR 24-DEC-1998; 98IN-DE003825.
XX
PA (COUL ) CSIR COUNCIL SCI IND RES.
XX
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
DR WPI; 2000-516032/47.
XX
PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
PS Disclosure; Fig 17b; 58pp; English.
XX
CC This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrid have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1541 BP; 497 A; 328 C; 335 G; 381 T; 0 U; 0 Other;

```

Query Match 71.4%; Score 1496.2; DB 3; Length 1541;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 588 ATTGCTGACCTGAGTGGCTCTAGACCGTCAATCTGTCAACAACAGCAATTTGTTGT 647
      |||
DB 33 ATAGCTGTCTCTGAATGGCTACTAGATGTCTCTTGTAAATTAACAGCAATTTGTTGT 92
QY 648 AGCGTTGCTGTAAGTCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 707
      |||
DB 93 AGCGTTGCTGTAAGTCTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 152
QY 708 GATCTAATCAACGACCTGCTCATGAGAGAAAGACAGGCAAGGCTTAAGTCAAAATCA 767
      |||
DB 153 GATCTAATCAACGACCTGCTCATGAGAGAAAGACAGGCAAGGCTTAAGTCAAAATCA 212
QY 768 AAACCATTTGCTAAGTATGAGGCGCGGATGTCATTAACCTTGAAGAAAGCTGACTTACTA 827
      |||
DB 213 AAACCATTTGCTAAGTATGAGGCGCGGATGTCATTAACCTTGAAGAAAGCTGACTTACTA 272

```

QY 828 AAGCTATTCAAGAACAATTGATGCTAACGTCACAGTAAACAAGACTTCTTGAGGTC 887
 Db 273 AAGGCTATTCAAGAACAATTGATGCTAACGTCACAGTAAACAAGACTTCTTGAGGTC 332
 QY 888 ATTGATTTTGCAGGCGATGCAACATTAAGTATGAAAACGGCAAGGTTACTTTGCTGAC 947
 Db 333 ATTGATTTTGCAGGCGATGCAACATTAAGTATGAAAACGGCAAGGTTACTTTGCTGAC 392
 QY 948 AAAGATGTTCCGGTAACTTTCGCAACCCGACCTGTCGAAGATTTTTCGTAAGGGAAT 1007
 Db 393 AAAGATGTTCCGGTAACTTTCGCAACCCGACCTGTCGAAGATTTTTCGTAAGGGAAT 452
 QY 1008 GTGCGGTTAGACCATTAATAAGAAAACAATACAAAAACCAAGGAAATCTGTTGATGTG 1067
 Db 453 GTGCGGTTAGACCATTAATAAGAAAACAATACAAAAACCAAGGAAATCTGTTGATGTG 512
 QY 1068 GAATATATCTGTACAGTTTACTCCTTAAACCTGATGACGATTTTCAACAGGTTCTCAA 1127
 Db 513 GAATATATCTGTACAGTTTACTCCTTAAACCTGATGACGATTTTCAACAGGTTCTCAA 572
 QY 1128 GATACCTAAGCTATTGAAGAAACATAGCATGCGTGAACAACATCAATCAATCTCAAGAAATTA 1187
 Db 573 GATACCTAAGCTATTGAAGAAACATAGCATGCGTGAACAACATCAATCAATCTCAAGAAATTA 632
 QY 1188 GCTCAAGCACAAGACATTTTAAACAAAACCAACAGGCTATACGATTTATGAAGTAC 1247
 Db 633 GCTCAAGCACAAGACATTTTAAACAAAACCAACAGGCTATACGATTTATGAAGTAC 692
 QY 1248 TCCCTCAATCGTCACTCAGTACATGACATTTTCCGTAAGATTTTACCAATGATCAAGAG 1307
 Db 693 TCCCTCAATCGTCACTCAGTACATGACATTTTCCGTAAGATTTTACCAATGATCAAGAG 752
 QY 1308 TTTAACCTAACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCTGCTG 1367
 Db 753 TTTAACCTAACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCTGCTG 812
 QY 1368 AATGAAGAAATAAACAACATGACCTGATCTGAGAAATTTAGTCTTAAAAAGGG 1427
 Db 813 AATGAAGAAATAAACAACATGACCTGATCTGAGAAATTTAGTCTTAAAAAGGG 872
 QY 1428 GAAAGCCGATGATCCCTTGTGATCGAGTCACTTGAACCTGTCACCATCAAAATAGCTT 1487
 Db 873 GAAAGCCGATGATCCCTTGTGATCGAGTCACTTGAACCTGTCACCATCAAAATAGCTT 932
 QY 1488 GATGTGATACCAACGAATGCTTAAAAAGTGAAGAGCTCTTAAACAGCTAGGGAACGTAC 1547
 Db 933 GATGTGATACCAACGAATGCTTAAAAAGTGAAGAGCTCTTAAACAGCTAGGGAACGTAC 992
 QY 1548 TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTTAACTCTACAAACAATCTC 1607
 Db 993 TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTTAACTCTACAAACAATCTC 1052
 QY 1608 GATGCTTTTGGATTATGACTATACCTTAACTGAAAAGTGAAGAGATTAATCAGATGAC 1667
 Db 1053 GATGCTTTTGGATTATGACTATACCTTAACTGAAAAGTGAAGAGATTAATCAGATGAC 1112
 QY 1668 ACCAACCGTATCATTAACCGTTTATATGAGGCAAGGCAACCGAAGAGAAATGCTAGCTAC 1727
 Db 1113 ACCAACCGTATCATTAACCGTTTATATGAGGCAAGGCAACCGAAGAGAAATGCTAGCTAC 1172
 QY 1728 CATTTAGCTGTGTGTGCGCCAGGCGCAACAATGTTACCATGCTGAGAAAGTGTGAT 1787
 Db 1173 CATTTAGCTGTGTGTGCGCCAGGCGCAACAATGTTACCATGCTGAGAAAGTGTGAT 1232
 QY 1788 CATGCTGTGGGACTTCTATGTGTGTGAGAAACGTTGGGAGAAACCCCTACCAAGGCTGG 1847
 Db 1233 CATGCTGTGGGACTTCTATGTGTGTGAGAAACGTTGGGAGAAACCCCTACCAAGGCTGG 1292
 QY 1848 ATGATGATGATTTGATCTTCCCTGAGGAGAGAGCGACGATCACTTGCACTTCTAGA 1907
 Db 1293 ATGATGATGATTTGATCTTCCCTGAGGAGAGAGCGACGATCACTTGCACTTCTAGA 1352

QY 1908 AATAGATCAACGATCAGGACACAAGACATCTTATAGAAATTGAGACACCTGAGCAAG 1967
 Db 1353 AATAGATCAACGATCAGGACACAAGACATCTTATAGAAATTGAGACACCTGAGCAAG 1412
 QY 1968 AAGATTAATCGAGAAACCTGCTCCAGTGCATCTGCAACGAGCAACGCGCCGAGAGATGCG 2027
 Db 1413 AAGATTAATCGAGAAACCTGCTCCAGTGCATCTGCAACGAGCAACGCGCCGAGAGATGCG 1472
 QY 2028 AAGTGTGAGAGGACACACCTGTCGAGAACCAATGAGGAGATCTGAGCCCTTCAACCAT 2087
 Db 1473 AAGTGTGAGAGGACACACCTGTCGAGAACCAATGAGGAGATCTGAGCCCTTCAACCAT 1532
 QY 2088 GTTCGTTAG 2096
 Db 1533 GTTCGTTAG 1541
 RESULT 4
 AAA37637
 ID AAA37637 standard; DNA; 1661 BP.
 XX
 AC AAA37637;
 XX
 DT 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE Chimeric SK-FBD coding sequence.
 XX
 KM Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KM plasminogen; human; fibrinectin; thrombolytic therapy;
 KM cardiovascular disorder; ss.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Homo sapiens.
 OS Chimeric.
 PN EP1024192-A2.
 PD 02-AUG-2000.
 XX
 PF 23-DEC-1999; 99EP-00310541.
 XX
 PR 24-DEC-1998; 98IN-DE003825.
 XX
 PA (COUL) CSIR COUNCIL SCT IND RES.
 XX
 PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX
 DR WPI; 2000-516032/47.
 XX
 PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX
 PS Example 5; Fig 19b; 58bp; English.
 XX
 CC This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins

CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1661 BP; 529 A; 370 C; 357 G; 405 T; 0 U; 0 Other;

Query Match 55.6%; Score 1165.8; DB 3; Length 1661;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

OY 588 ATTGCTGACCTGAGTGGCTGCTGACCGTCACTGTCACCAACAGCAATGGTGT 647
    |||||
DB 183 ATTGCTGACCTGAGTGGCTGCTGACCGTCACTGTCACCAACAGCAATGGTGT 242
    |||||
OY 648 AGCGTTGCTGTACTGTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGAATC 707
    |||||
DB 243 AGCGTTGCTGTACTGTGAGGGGACGATCAAGACATTAAGTCTTAAATTTTGAATC 302
    |||||
OY 708 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
    |||||
DB 303 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 362
    |||||
OY 768 AAACCATTTGCTAGTATAGGGGCGGATGTCATATAAATTGGAAGGCTTAAGTCTA 827
    |||||
DB 363 AAACCATTTGCTAGTATAGGGGCGGATGTCATATAAATTGGAAGGCTTAAGTCTA 422
    |||||
OY 828 AAGGCTATTCAGAAACATTAAGTCTTAAGTCCAGATTAAGCACTACTTGAAGTTC 887
    |||||
DB 423 AAGGCTATTCAGAAACATTAAGTCTTAAGTCCAGATTAAGCACTACTTGAAGTTC 482
    |||||
OY 888 ATTGATTTTTCAGAACGATGCAACCATTAAGTGAAGGCAAGGCTTACTTGTGAC 947
    |||||
DB 483 ATTGATTTTTCAGAACGATGCAACCATTAAGTGAAGGCAAGGCTTACTTGTGAC 542
    |||||
OY 948 AAAGATGTTGGGTAACCTTGGCGACCCCAACCTGTCACAAATTTTGTCTAAGGGACAT 1007
    |||||
DB 543 AAAGATGTTGGGTAACCTTGGCGACCCCAACCTGTCACAAATTTTGTCTAAGGGACAT 602
    |||||
OY 1008 GTGCGCGTTAGACCATATAAGAAAAACAATACAAACCAAGCAATCTGTGATGTG 1067
    |||||
DB 603 GTGCGCGTTAGACCATATAAGAAAAACAATACAAACCAAGCAATCTGTGATGTG 662
    |||||
OY 1068 GAATATACCTGTAAGTTACTCCCTTAAACCTGTATGACGATTTCAAGCAGGTCTCAA 1127
    |||||
DB 663 GAATATACCTGTAAGTTACTCCCTTAAACCTGTATGACGATTTCAAGCAGGTCTCAA 722
    |||||
OY 1128 GATACCTAAGCTATTGAAAAACATAGCTATGCTGACACATCAATCTCAAGAAATTACTA 1187
    |||||
DB 723 GATACCTAAGCTATTGAAAAACATAGCTATGCTGACACATCAATCTCAAGAAATTACTA 782
    |||||
OY 1188 GCTCAAGCAACAAGCATTTTAAACAAAAACACCAGGCTATAGATTTTATGAACGTGAC 1247
    |||||
DB 783 GCTCAAGCAACAAGCATTTTAAACAAAAACACCAGGCTATAGATTTTATGAACGTGAC 842
    |||||
OY 1248 TCCTCAATGCTCACTCATGACATGACATTTCCGTACGANTTTTACCAATGATCAAGAG 1307
    |||||
DB 843 TCCTCAATGCTCACTCATGACATGACATTTCCGTACGANTTTTACCAATGATCAAGAG 902
    |||||
OY 1308 TTTACTTACCGTGTAAATAATTCGGGAAACAGCTTATAGATCAATATAAATCTGTGTG 1367
    |||||
DB 903 TTTACTTACCGTGTAAATAATTCGGGAAACAGCTTATAGATCAATATAAATCTGTGTG 962
    |||||
OY 1368 AATGAAGAAATAAACAACACGACCTGATCTGAGAAATTTTACGTCCTTAAAAAGGG 1427
    |||||
DB 963 AATGAAGAAATAAACAACACGACCTGATCTGAGAAATTTTACGTCCTTAAAAAGGG 1022
    |||||
OY 1428 GAAAGACCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCACCATCAAAATAGCTT 1487
    |||||
DB 1023 GAAAGACCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCACCATCAAAATAGCTT 1082
    |||||

```

```

OY 1488 GATGCTGATACCAACGAATTCCTAATAAAGTAGACAGCTCTTAACAGTAGCAAGTAC 1547
    |||||
DB 1083 GATGCTGATACCAACGAATTCCTAATAAAGTAGACAGCTCTTAACAGTAGCAAGTAC 1142
    |||||
OY 1548 TTAGACTTCAGAGATTTATACGATCTCTGTATTAAGGCTTAACCTACTACCAATCTTC 1607
    |||||
DB 1143 TTAGACTTCAGAGATTTATACGATCTCTGTATTAAGGCTTAACCTACTACCAATCTTC 1202
    |||||
OY 1608 GATGCTTTTGTATTAATGACCTATACCTTAACCTGGAAGAGATTAATACGATGAC 1667
    |||||
DB 1203 GATGCTTTTGTATTAATGACCTATACCTTAACCTGGAAGAGATTAATACGATGAC 1262
    |||||
OY 1668 ACCAACCGTATCATACCGTTTATATGAGGCAAGCAGACCGAAGAGAGATCTAGTAC 1727
    |||||
DB 1263 ACCAACCGTATCATACCGTTTATATGAGGCAAGCAGACCGAAGAGAGATCTAGTAC 1322
    |||||
OY 1728 CATTTAGCTGTGTGTCGCGCAGGCGCAACAGATTGTAC 1764
    |||||
DB 1323 CATTTAGCTGTGTGTCGCGCAGGCGCAACAGATTGTTC 1359
    |||||

```

RESULT 5

AAK80497
 ID AAK80497 standard; cDNA; 2385 BP.

XX AAK80497;

XX 26-AUG-1999 (first entry)

DE Streptokinase and maltose binding protein fusion protein encoding cDNA.

XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;

KW nSK; Bacterial; blood clot; thrombotic condition;

KW myocardial infarction; venous thrombosis; pulmonary embolism;

XX cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

OS Synthetic.

XX WO9331247-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-US026694.

XX 15-DEC-1997; 97US-0069497P.

XX (HARD) HARVARD COLLEGE.

XX Read GLI;

XX WPI; 1999-395183/33.

XX P-PSDB; AAY24797.

XX N-terminally deleted streptokinase.

XX Example; Page 45-48; 73pp; English.

XX The present invention describes an isolated bacterial protein that

CC induces fibrin-dependent plasminogen activation in a pharmacological

CC composition for dissolving blood clots. Also described are: (1) a

CC composition comprising an isolated modified streptokinase; the

CC modification being removal of amino acid residues in the amino terminus;

CC (2) a method for dissolving a blood clot in a subject, comprising

CC administering to the subject a fibrin-dependent streptokinase protein; a

CC nucleic acid (I) encoding a modified bacterial streptokinase; (3) an

CC expression vector comprising (I); and (4) a host cell transformed with

CC the expression vector of (3). The pharmaceutical composition comprising a

CC bacterial fibrin-dependent plasminogen activator is useful for dissolving

CC blood clots in patients with a thrombotic condition, e.g. myocardial

CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,

CC graft thrombosis and arterial thrombosis. The modified streptokinase can

CC also be used in non-human mammals. Streptokinase activation of

CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for streptolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC encodes a streptokinase and maltose binding protein fusion protein from
 CC an example of the present invention

XX Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;

Query Match 54.9%; Score 1150.8; DB 2; Length 2385;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1158; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

QY 573 TTCACCGATGTCGTATTTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAC 632
DB 1129 TCCATCGAGGTAGGATTTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAC 1188
QY 633 AGCCAAATGGTTGTTGACGTTGCTGTACTGTGAGGGGACGAATCAAGACATTAGTCTT 692
DB 1189 AGCCAAATGGTTGTTGACGTTGCTGTACTGTGAGGGGACGAATCAAGACATTAGTCTT 1248
QY 693 AAATTTTGGAAATCGATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGCC 752
DB 1249 AAATTTTGGAAATCGATCTAACATCAGACCTGCTCATGAGGAAAGACAGAGCAAGCC 1308
QY 753 TTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGGGGAGTGCATATACTTAG 812
DB 1309 TTAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGGGGAGTGCATATACTTAG 1368
QY 813 AAAGCTGACTTACTAAAGGCTATTCAAGAACATTTGCTTAAAGTCAAGTAAAGCAG 872
DB 1369 AAAGCTGACTTACTAAAGGCTATTCAAGAACATTTGCTTAAAGTCAAGTAAAGCAG 1428
QY 873 GACTACTTGAAGTCAATGATTTTTCGACGATGCAACCAATTAAGTGAACCGCAAG 932
DB 1429 GACTACTTGAAGTCAATGATTTTTCGACGATGCAACCAATTAAGTGAACCGCAAG 1488
QY 933 GTTACTTTTGTGCAAGATGCTTGGGTAACTTTGCCGACCACTGTCCAGAAATTT 992
DB 1489 GTTACTTTTGTGCAAGATGCTTGGGTAACTTTGCCGACCACTGTCCAGAAATTT 1548
QY 993 TTGCTAAGCGGACATGTCGGCTTGAACCATATAAGAAACCAATCAAAACCAAGCG 1052
DB 1549 TTGCTAAGCGGACATGTCGGCTTGAACCATATAAGAAACCAATCAAAACCAAGCG 1608
QY 1053 AAATCTGTTGATGGAATATATCTGTACAGTTTATCTCCCTTAAACCTGATGACGATTT 1112
DB 1609 AAATCTGTTGATGGAATATATCTGTACAGTTTATCTCCCTTAAACCTGATGACGATTT 1668
QY 1113 AGACCAAGTCTCAAGATTAAGCTATTGAAAACATGACTATCGGTGACACCATTCACA 1172
DB 1669 AGACCAAGTCTCAAGATTAAGCTATTGAAAACATGACTATCGGTGACACCATTCACA 1728
QY 1173 TCTCAAGAAATTAAGTCAAGCACAAGCAATTTTAAACAAACCAACCCGCGTAAACG 1232
DB 1729 TCTCAAGAAATTAAGTCAAGCACAAGCAATTTTAAACAAACCAACCCGCGTAAACG 1788
QY 1233 AATTATGAACGTGACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTA 1292
DB 1789 AATTATGAACGTGACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTA 1848
QY 1293 CCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAAT 1352
DB 1849 CCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAAT 1908
QY 1353 AAAAAATCTGCTGATGAAGAATAAACAACCTGACCTGATCTTGAAGAAATATTAC 1412
DB 1909 AAAAAATCTGCTGATGAAGAATAAACAACCTGACCTGATCTTGAAGAAATATTAC 1968
QY 1413 GTCTTAAAAAAGGGGAAAAAGCGTATGATCCCTTTGATGCGACGACTTGAACCTGTT 1472
DB 1969 GTCTTAAAAAAGGGGAAAAAGCGTATGATCCCTTTGATGCGACGACTTGAACCTGTT 2028

```

```

QY 1473 ACCATCAATATAGTTGATGTCGATACCAAGATTTGCTAAAAAGTAGAGCTCTTAA 1532
DB 2029 ACCATCAATATAGTTGATGTCGATACCAAGATTTGCTAAAAAGTAGAGCTCTTAA 2088
QY 1533 GCTAGCGAAAGTAACTTGAAGTTCAGAGATTTTAAAGATCTCTGTATAGGCTTAA 1592
DB 2089 GCTAGCGAAAGTAACTTGAAGTTCAGAGATTTTAAAGATCTCTGTATAGGCTTAA 2148
QY 1593 CTCTCAACAATCTGAGTCTTTGGTATATAGACTATACCTTAACCTGAAAAAGTAGAG 1652
DB 2149 CTCTCAACAATCTGAGTCTTTGGTATATAGACTATACCTTAACCTGAAAAAGTAGAG 2208
QY 1653 GATAATCAGATGACACCAACCGTATCATPAACGTTTATATAGGCAAGGACCCGAAGA 1712
DB 2209 GATAATCAGATGACACCAACCGTATCATPAACGTTTATATAGGCAAGGACCCGAAGA 2268
QY 1713 GAGAAATGCTAGCTTACATTTAGCTGTGTGT 1742
DB 2269 GAGAAATGCTAGCTTACATTTAGCTGTGTGT 2298

```

RESULT 6

AAA37622
 ID AAA37622 standard; DNA; 1377 BP.

XX AAA37622;

AC 15-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)

XX Streptokinase-NTN gene.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

PN EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99BP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) GSIR COUNCIL SCI IND RES.

PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

PI Yadav M;

DR WPI; 2000-516032/47.

PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibronectin.

PS Example 1; Fig 11; 58pp; English.

CC This sequence represents a streptokinase-NTN (SK-NTN) gene (where NTN
 CC stands for N-terminally repaired with native sequence). The invention
 CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
 CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
 CC activation, and fibrin binding regions of human fibronectin, which are
 CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
 CC possesses the ability to bind with fibrin independently and also
 CC characteristically retains a PG activation ability which becomes evident
 CC only after a pronounced duration, or lag, after exposure of the PA to a
 CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
 CC domain polypeptides are useful in thrombolytic therapy for various kinds
 CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
 CC as well as kinetics of plasminogen activation that are distinct from that

CC of natural streptokinase in being characterised by a temporary delay, or
 CC lag of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;

Query Match 54.7%; Score 1147; DB 3; Length 1377;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 588 ATTGCTGACCTGAGTGGCTGCTGACCGCTCCATCTGTCAACAAAGCAATGGTTGT 647
DB 133 ATTGCTGACCTGAGTGGCTGCTGACCGCTCCATCTGTCAACAAAGCAATGGTTGT 192
QY 648 AGCGTTGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGCTTAATTTTGAATC 707
DB 193 AGCGTTGCTGCTGCTGCTGAGGGGACGAATCAAGACATTAGCTTAATTTTGAATC 252
QY 708 GATCTAATCAGACGCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 767
DB 253 GATCTAATCAGACGCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 312
QY 768 AAACCAATTTGCTACGTAGTAGTGGCGCGATGTCACATTAATCTTGAGAAAGCTTA 827
DB 313 AAACCAATTTGCTACGTAGTAGTGGCGCGATGTCACATTAATCTTGAGAAAGCTTA 372
QY 828 AAGGCTATTCAAAACAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
DB 373 AAGGCTATTCAAAACAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
QY 888 ATTGATTTTTCAGCGCATGCAACATTAATGATGCAAGGCAAGGCTTAAGTCTGAC 947
DB 433 ATTGATTTTTCAGCGCATGCAACATTAATGATGCAAGGCAAGGCTTAAGTCTGAC 492
QY 948 AAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1007
DB 493 AAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
QY 1008 GTGCGGCTTGAACCATTAATTAAGAAACCAATTAATTAAGAAACCAATTAATTA 1067
DB 553 GTGCGGCTTGAACCATTAATTAAGAAACCAATTAATTAAGAAACCAATTAATTA 612
QY 1068 GAATATACGTAGATTTTACTCCCTTAACCTGATGACATTTTCAACCAAGGCTCAA 1127
DB 613 GAATATACGTAGATTTTACTCCCTTAACCTGATGACATTTTCAACCAAGGCTCAA 672
QY 1128 GATCTAAGCTATTGAAAACATAGCTATGCTGACACCATCATCTCAAGATTTACTA 1187
DB 673 GATCTAAGCTATTGAAAACATAGCTATGCTGACACCATCATCTCAAGATTTACTA 732
QY 1188 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTTAATGATTTAAGAG 1247
DB 733 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTTAATGATTTAAGAG 792
QY 1248 TCCTCAATGCTCATGATGACATTAATTTTCCGTACGATTTTCAACATGATGACAG 1307
DB 793 TCCTCAATGCTCATGATGACATTAATTTTCCGTACGATTTTCAACATGATGACAG 852
QY 1308 TTTACTTACCGTGTAAATCGGGAACAGCTTAATGATCAATTAATAATCTGCTG 1367
DB 853 TTTACTTACCGTGTAAATCGGGAACAGCTTAATGATCAATTAATAATCTGCTG 912
QY 1368 AATGAAGAAATAAACAACACTGACCTGATCTGAGAAATATTAACGCTTTAAATAAG 1427
DB 913 AATGAAGAAATAAACAACACTGACCTGATCTGAGAAATATTAACGCTTTAAATAAG 972
QY 1428 GAAAGACCGATATCCCTTTGATCGAGTCACTTGAACCTGTCACCAATCAATACGTT 1487

```

```

DB 973 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTCACCATCAATACGTT 1032
QY 1488 GATTCGATATCAACAGAAATTTCTTAATAAGAGAGAGCTTTAAACAGCTAGGAACTAC 1547
DB 1033 GATTCGATATCAACAGAAATTTCTTAATAAGAGAGAGCTTTAAACAGCTAGGAACTAC 1092
QY 1548 TTAGACTTCAGAGATTTATATACATCTCGATTAAGGCTTAATCTTACAAATCTC 1607
DB 1093 TTAGACTTCAGAGATTTATATACATCTCGATTAAGGCTTAATCTTACAAATCTC 1152
QY 1608 GATCCTTTTGGTATTAATGACCTAATACCTTAATGAAAAGTAAAGATTAATCAGATGAC 1667
DB 1153 GATCCTTTTGGTATTAATGACCTAATACCTTAATGAAAAGTAAAGATTAATCAGATGAC 1212
QY 1668 ACCAAGCGTATCATTAACCGTTTAAATGAGGCAAGGACCCGAAGAGAGAAATGCTAC 1727
DB 1213 ACCAAGCGTATCATTAACCGTTTAAATGAGGCAAGGACCCGAAGAGAGAAATGCTAT 1272
QY 1728 CATTAGCTGCTGCT 1742
DB 1273 CATTAGCTATGAT 1287

RESULT 7
AAK80492
ID AAK80492 standard; cDNA; 1242 BP.
XX
AC AAK80492;
XX
DT 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
DE Streptococcus equisimilis native streptokinase encoding cDNA.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nsk; RSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN W09931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US026694.
XX
PR 15-DEC-1997; 97US-0069497P.
XX
PA (HARD ) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1999-395183/33.
XX
DR P-PSDB; AAY24794.
XX
N-terminally deleted streptokinase.
XX
PS Claim 44; Page 58-60; 73pp; English.
XX
CC The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (II) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (I); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a
CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC blood clots in patients with a thrombotic condition, e.g. myocardial

```

CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC encodes native streptokinase (nsK). (Updated on 17-OCT-2003 to
 CC standardise OS field)

XX Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;

Query Match 54.6%; Score 1145.4; DB 2; Length 1242;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGGCTGCTGACCGCTCATCTGTCAACACAGCCAAATGGTGTG 647
 DB 1 ATTGCTGACCTGAGTGGCTGCTGACCGCTCATCTGTCAACACAGCCAAATGGTGTG 60
 QY 648 AGCGTTCGTGTAAGTGGGAGCGAATCAAGCACTTACTTAAATTTTGAATC 707
 DB 61 AGCGTTCGTGTAAGTGGGAGCGAATCAAGCACTTACTTAAATTTTGAATC 120
 QY 708 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 767
 DB 121 GATCTAACATCAGACCTGCTCATGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 180
 QY 768 AAACCACTTGTCTAGTATGCGCGCATGTCAATAACTTGAAGAAAGTGAATCTACTA 827
 DB 181 AAACCACTTGTCTAGTATGCGCGCATGTCAATAACTTGAAGAAAGTGAATCTACTA 240
 QY 828 AAGGCTATTCAGAAACAATTGATCGCTAAGCTCCAGTACAGAGCTACTTTGAAGTC 887
 DB 241 AAGGCTATTCAGAAACAATTGATCGCTAAGCTCCAGTACAGAGCTACTTTGAAGTC 300
 QY 888 ATTGATTTTTCAGACGATGCAACCATTAAGTGAAGACGCAAGTCTACTTTGCTGAC 947
 DB 301 ATTGATTTTTCAGACGATGCAACCATTAAGTGAAGACGCAAGTCTACTTTGCTGAC 360
 QY 948 AAGATGTTGCTGTAACCTTGGCGACCCAACTGTCCAGAAATTTTGTAAAGCGGACAT 1007
 DB 361 AAGATGTTGCTGTAACCTTGGCGACCCAACTGTCCAGAAATTTTGTAAAGCGGACAT 420
 QY 1008 GTGCGGCTTACACATTTAAAGAAAAACAATCAAAACCAAGGAAATCTGTGATGTG 1067
 DB 421 GTGCGGCTTACACATTTAAAGAAAAACAATCAAAACCAAGGAAATCTGTGATGTG 480
 QY 1068 GAATATACGTATAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCCAGGCTCAAA 1127
 DB 481 GAATATACGTATAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCCAGGCTCAAA 540
 QY 1128 GATATACGTATTTGAAAAACATAGCTATGCTGACACATCAATCAAGATTTACTA 1187
 DB 541 GATATACGTATTTGAAAAACATAGCTATGCTGACACATCAATCAAGATTTACTA 600
 QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATAGATTTTGAAGCGTAC 1247
 DB 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGGCTATAGATTTTGAAGCGTAC 660
 QY 1248 TCCTCAATGCTGATGATGACATGACATTTTCCGTACGATTTTCAAGATGACAGAG 1307
 DB 661 TCCTCAATGCTGATGATGACATGACATTTTCCGTACGATTTTCAAGATGACAGAG 720
 QY 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATTAATAATCGTGTG 1367
 DB 721 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATTAATAATCGTGTG 780
 QY 1368 AATGAAGAAATTAACAACATGACCTGATCTGAGAAATATTACGTCCTTTAAAAAGGG 1427
 DB 781 AATGAAGAAATTAACAACATGACCTGATCTGAGAAATATTACGTCCTTTAAAAAGGG 840

QY 1428 GAAAAACCGTATATCCCTTTGATGCGACGTCACTTGAAACTGTTCAACATCAATACGTT 1487
 DB 841 GAAAAACCGTATATATCCCTTTGATGCGACGTCACTTGAAACTGTTCAACATCAATACGTT 900
 QY 1488 GATGTCGATACCAAGAAATTTGCTTAAAAAGTGAAGACGCTTTAAACAGTACGGAACGTAAC 1547
 DB 901 GATGTCGATACCAAGAAATTTGCTTAAAAAGTGAAGACGCTTTAAACAGTACGGAACGTAAC 960
 QY 1548 TTAGACTTACAGATTTTATAGATCCCTGATGATGAGGCTTAACTACTTCACAAATATTC 1607
 DB 961 TTAGACTTACAGATTTTATAGATCCCTGATGATGAGGCTTAACTACTTCACAAATATTC 1020
 QY 1608 GATGCTTTTGTATTTATGACATTAACCTTAAGTGAAGAAAGTGAAGATATACAGATGAC 1667
 DB 1021 GATGCTTTTGTATTTATGACATTAACCTTAAGTGAAGAAAGTGAAGATATACAGATGAC 1080
 QY 1668 ACCAACCGTATATTAACGCTTTATATGAGCAAGCAAGGAGAGAGATGCTAGCTAC 1727
 DB 1081 ACCAACCGTATATTAACGCTTTATATGAGCAAGCAAGGAGAGAGATGCTAGCTAT 1140
 QY 1728 CATTTAGCTGTGCT 1742
 DB 1141 CATTTAGCTGTATGAT 1155

RESULT 8

AAA37633
 ID AAA37633 standard; DNA; 1245 BP.

AAA37633;

AC 15-SEP-2003 (revised)

DT 13-OCT-2000 (first entry)

XX S. equisimilis streptokinase coding sequence.

DE Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers

FT CDS 1..1245

FT /tag= a
 FT /product= "streptokinase"

PN EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUNL) CSIR COUNCIL SCT IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;

XX WPI; 2000-516032/47.

XX P-PSDB; AAY90282.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.

XX Example 3; Fig 3; 58pp; English.

XX

XX This sequence represents the human Streptococcus equisimilis
 CC streptokinase coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and

CC fibrin binding regions of human fibronectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a pg activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human Pg. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolytic). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-Sep-2003 to standardise OS field)

CC Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match 54.6%; Score 1145.4; DB 3; Length 1245;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTGGTTG 647
 Db 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCAATTGGTTG 60
 Qy 648 AGCGTTGCTGTAAGTGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 707
 Db 61 AGCGTTGCTGTAAGTGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 120
 Qy 708 GATCTAACATCAGACCTGCTCATGAGAAACAGAGCAAGGCTTAAGTCCAAATCA 767
 Db 121 GATCTAACATCAGACCTGCTCATGAGAAACAGAGCAAGGCTTAAGTCCAAATCA 180
 Qy 768 AAACCATTTGCTACTGATGAGGAGGAGATGCTCACTAACTTGAGAAAGCTTACTA 827
 Db 181 AAACCATTTGCTACTGATGAGGAGGAGATGCTCACTAACTTGAGAAAGCTTACTA 240
 Qy 828 AAGGCTATTCAGAAACATTCATGCTCAAGTCAAGACGACTACTTGAAGTTC 887
 Db 241 AAGGCTATTCAGAAACATTCATGCTCAAGTCAAGACGACTACTTGAAGTTC 300
 Qy 888 ATTGATTTTGCAGACGATGCAACATTAAGTCAAGAAAGGCTTACTTGTGAC 947
 Db 301 ATTGATTTTGCAGACGATGCAACATTAAGTCAAGAAAGGCTTACTTGTGAC 360
 Qy 948 AAAGATGCTGGGTAACCTTGGCCGACCACTGCTCAAGAAATTTTGTAGGGAAT 1007
 Db 361 AAAGATGCTGGGTAACCTTGGCCGACCACTGCTCAAGAAATTTTGTAGGGAAT 420
 Qy 1008 GTGCGGTTAGACCATATAAGAAAAACAATCAAAACCAAGGAATCTGTGATGTG 1067
 Db 421 GTGCGGTTAGACCATATAAGAAAAACAATCAAAACCAAGGAATCTGTGATGTG 480
 Qy 1068 GAATATAGCTGATGATTTTCTCCCTTAAACCTGATGAGATTTCAAGACAGTCTCAA 1127
 Db 481 GAATATAGCTGATGATTTTCTCCCTTAAACCTGATGAGATTTCAAGACAGTCTCAA 540
 Qy 1128 GATACCTAAGCTATTTGAAAAACATAGCTATGAGTCAACATCTCAAGAAATTTCTA 1187
 Db 541 GATACCTAAGCTATTTGAAAAACATAGCTATGAGTCAACATCTCAAGAAATTTCTA 600
 Qy 1188 GCTCAAGCAACAAAGATTTTAAACAAACACACAGGCTATAGATTTTGAAGTGAAC 1247
 Db 601 GCTCAAGCAACAAAGATTTTAAACAAACACACAGGCTATAGATTTTGAAGTGAAC 660
 Qy 1248 TCCTCAATGCTCACTCATGACCAATGACATTTTCGTAAGATTTTACCAATGATCAAGAG 1307
 Db 661 TCCTCAATGCTCACTCATGACCAATGACATTTTCGTAAGATTTTACCAATGATCAAGAG 720

Qy 1308 TTACTTACCGTGTAAATAATCGGAAACACAGCTTATAGATCAATATAAAAATCTGCTG 1367
 Db 721 TTACTTACCGTGTAAATAATCGGAAACACAGCTTATAGATCAATATAAAAATCTGCTG 780
 Qy 1368 AATGAAGAAATTAACACACAGCTGATCTCTGGAATATTTAGTCTTAAAAAAGG 1427
 Db 781 AATGAAGAAATTAACACACAGCTGATCTCTGGAATATTTAGTCTTAAAAAAGG 840
 Qy 1428 GAAAAGCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTCACCATCAATAGCTT 1487
 Db 841 GAAAAGCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTCACCATCAATAGCTT 900
 Qy 1488 GATGCTGATACCAAGAAATTCCTAAAAAGTGAAGAGCTTTAAACAGTGAAGAGTAAAC 1547
 Db 901 GATGCTGATACCAAGAAATTCCTAAAAAGTGAAGAGCTTTAAACAGTGAAGAGTAAAC 960
 Qy 1548 TTAGACTTCAGAGATTTATATAGATCTCTGATTAAGGCTAAACTCTTACCAATCTC 1607
 Db 961 TTAGACTTCAGAGATTTATATAGATCTCTGATTAAGGCTAAACTCTTACCAATCTC 1020
 Qy 1608 GATGCTTTGGTATATAGACTATATACCTTAACTGAAAAGTGAAGATTAACAGATGAC 1667
 Db 1021 GATGCTTTGGTATATAGACTATATACCTTAACTGAAAAGTGAAGATTAACAGATGAC 1080
 Qy 1668 ACCAAGCTATCATACCGTTTATATGAGCAAGCAACCGAAGAGAGAAATGCTAGTAC 1727
 Db 1081 ACCAAGCTATCATACCGTTTATATGAGCAAGCAACCGAAGAGAGAAATGCTAGTAC 1140
 Qy 1728 CATTAGCTGCTGCT 1742
 Db 1141 CATTAGCTGCTGCT 1155

RESULT 9

ABA05546
 ID ABA05546 standard; cDNA; 1254 BP.

XX ABA05546;

XX 26-FEB-2002 (first entry)

XX Streptokinase cDNA.

XX Streptokinase; cerebroprotective; cardiact; gene therapy; fusion;

XX Vaso dilator; thrombolytic; angina; myocardial infarction; stroke;

XX gene therapy; maxadilan; ss.

XX Unidentified.

XX W0200185100-A2.

XX 15-NOV-2001.

XX 10-MAY-2001; 2001WO-US015209.

XX 11-MAY-2000; 2000US-00569920.

XX (GEHO) GEN HOSPITAL CORP.

XX Reddy VB, Lerner B;

XX WPI; 2002-062184/08.

XX New fusion protein or conjugate, useful for treating unstable angina,

XX acute myocardial infarction or stroke, comprises a vaso dilator

XX polypeptide and a thrombolytic polypeptide, or active fragments of the

XX polypeptides.
 XX Example 1; Fig 2; 37pp; English.
 XX The invention relates to a fusion protein or a conjugate comprising a
 CC vaso dilator polypeptide, or its active fragment, and a thrombolytic
 CC polypeptide or its active fragment. The protein is useful for treating a

CC subject suffering from a partially or totally occluded blood vessel,
 CC causing unstable angina, acute myocardial infarction or stroke. The
 CC polynucleotide encoding the polypeptide is useful in gene therapy. The
 CC vasodilatory action of the protein allows for the use of lower doses of a
 CC thrombolytic while maintaining the clot dissolving effectiveness of the
 CC thrombolytic, and the use of lower doses of the thrombolytic reduces
 CC associated side effects. The present sequence is the streptokinase cDNA
 CC used in the construction of a Maxadilan-Streptokinase fusion protein.
 CC Maxadilan is a vasodilator peptide produced by the salivary gland of the
 CC New World sand fly

XX Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;

Query Match 54.6%; Score 1145.4; DB 6; Length 1254;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCAATTGGTGT 647
DB 7 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCAATTGGTGT 66
QY 648 AGCGTGTGCTGCTGCTGCTGAGGGGAGCAATCAAGACTTAAATTTTGAATC 707
DB 67 AGCGTGTGCTGCTGCTGCTGAGGGGAGCAATCAAGACTTAAATTTTGAATC 126
QY 708 GATCTAATCAATCAAGACTGCTCAATGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 767
DB 127 GATCTAATCAATCAAGACTGCTCAATGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 186
QY 768 AAACCAATTTGCTAATGATGAGGCGCATGTCATTAATCTTGAAGAGTGAATCTA 827
DB 187 AAACCAATTTGCTAATGATGAGGCGCATGTCATTAATCTTGAAGAGTGAATCTA 246
QY 828 AAGGCTAATTCAGAAACAATGATGCTTAAGTCAAGTCAAGCACTACTTGAAGTTC 887
DB 247 AAGGCTAATTCAGAAACAATGATGCTTAAGTCAAGTCAAGCACTACTTGAAGTTC 306
QY 888 ATTGATTTTGGCAAGCAGATGCAACATTAATGATGCAAGGCAAGTCTTGTGCTGAC 947
DB 307 ATTGATTTTGGCAAGCAGATGCAACATTAATGATGCAAGGCAAGTCTTGTGCTGAC 366
QY 948 AAAGATGCTTCGGTAACTTCCGACCAACTCTGTCAGAAATTTTGTGAAGCGGACAT 1007
DB 367 AAAGATGCTTCGGTAACTTCCGACCAACTCTGTCAGAAATTTTGTGAAGCGGACAT 426
QY 1008 GTGCGCGTTAGACCATATTAAGAAAAACAATACAAACCAAGGAAATCTGTGATGTG 1067
DB 427 GTGCGCGTTAGACCATATTAAGAAAAACAATACAAACCAAGGAAATCTGTGATGTG 486
QY 1068 GAATATACCTGTACAGTTTAACTCCCTTAAACCTGATGACGATTTTCAACAGGCTCAAA 1127
DB 487 GAATATACCTGTACAGTTTAACTCCCTTAAACCTGATGACGATTTTCAACAGGCTCAAA 546
QY 1128 GATCTAATGCTATTTGAAGAACTAGTATGCTGAGACCAATCATCTCAAGAAATCTA 1187
DB 547 GATCTAATGCTATTTGAAGAACTAGTATGCTGAGACCAATCATCTCAAGAAATCTA 606
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCAAGGCTTAAGATTTAAGACGTGAC 1247
DB 607 GCTCAAGCAAAAGCAATTTTAAACAAACCAACCAAGGCTTAAGATTTAAGACGTGAC 666
QY 1248 TCTCTCAATGCTCACTGACATGACATGATTTTCCGTAAGATTTTCAAGATGAAAGAG 1307
DB 667 TCTCTCAATGCTCACTGACATGACATGATTTTCCGTAAGATTTTCAAGATGAAAGAG 726
QY 1308 TTACTTACCGTGTAAATAATGGGGAACAAGCTTAAGTCAATTAATAATAATCGTGTG 1367
DB 727 TTACTTACCGTGTGTAAATAATGGGGAACAAGCTTAAGTCAATTAATAATAATCGTGTG 786
QY 1368 AATGAAGAATAAACAACACTGACTGATCTGAGAAATATTAAGCTTTAAATAAAGGG 1427
DB 787 AATGAAGAATAAACAACACTGACTGATCTGAGAAATATTAAGCTTTAAATAAAGGG 846

```

```

QY 1428 GAAAAGCCGTATGATCCCTTGTAGTCAGTCACTTGAACATGCTGCACATCAAAATACGTT 1487
DB 847 GAAAAGCCGTATGATCCCTTGTAGTCAGTCACTTGAACATGCTGCACATCAAAATACGTT 906
QY 1488 GATGCTGATACCAAGAAATTTGTTAAAGTGAAGCAAGCTTTTAAACAGCTTAGCAAGTAAAC 1547
DB 907 GATGCTGATACCAAGAAATTTGTTAAAGTGAAGCAAGCTTTTAAACAGCTTAGCAAGTAAAC 966
QY 1548 TTAGACTTCAGAGATTTTATGAGTCTGCTGTATTAAGGCTTAAACTCTCTACAAACAATCTC 1607
DB 967 TTAGACTTCAGAGATTTTATGAGTCTGCTGTATTAAGGCTTAAACTCTCTACAAACAATCTC 1026
QY 1608 GATGCTTTTGGTATTTATGAGTCACTTAACTGGAAGAAAGTAGAGATTAATCAAGATGAC 1667
DB 1027 GATGCTTTTGGTATTTATGAGTCACTTAACTGGAAGAAAGTAGAGATTAATCAAGATGAC 1086
QY 1668 ACCAACCGTATCATACCGTTTATATGAGGCAAGCAACCGAAGAGAGAAATGCTGATAC 1727
DB 1087 ACCAACCGTATCATACCGTTTATATGAGGCAAGCAACCGAAGAGAGAAATGCTGATAC 1146
QY 1728 CATTTAGCTGTGCT 1742
DB 1147 CATTTAGCTGTATGAT 1161

```

RESULT 10

ABAO5547 standard; DNA; 8893 BP.

ABAO5547;

26-FEB-2002 (first entry)

Maxadilan-streptokinase fusion protein plasmid pTYB3maxsek.

Streptokinase; cerebroprotective; cardiac; gene therapy; fusion;

vasodilator; thrombolytic; angina; myocardial infarction; stroke;

gene therapy; maxadilan, sand fly; plasmid; de.

Lutzomyia longipalpis.

Undifferentiated.

Synthetic.

Chimeric.

WO200185100-A2.

15-NOV-2001.

10-MAY-2001; 2001WO-US015209.

11-MAY-2000; 2000US-00569920.

(GENO) GEN HOSPITAL CORP.

Reddy VB, Lerner E;

WPI; 2002-062184/08.

New fusion protein or conjugate, useful for treating unstable angina,

acute myocardial infarction or stroke, comprises a vasodilator

polypeptide and a thrombolytic polypeptide, or active fragments of the

polypeptides.

Example 1; Fig 3; 37pp; English.

The invention relates to a fusion protein or a conjugate comprising a

vasodilator polypeptide, or its active fragment, and a thrombolytic

polypeptide or its active fragment. The protein is useful for treating a

subject suffering from a partially or totally occluded blood vessel,

causing unstable angina, acute myocardial infarction or stroke. The

polynucleotide encoding the polypeptide is useful in gene therapy. The

vasodilatory action of the protein allows for the use of lower doses of a

thrombolytic while maintaining the clot dissolving effectiveness of the

CC thrombolytic, and the use of lower doses of the thrombolytic reduces
CC associated side effects. The present sequence is the plasmid used for the
CC preparation of Maxadilan-Streptokinase fusion protein. It contains
CC maxadilan cDNA, which encodes a vasodilator peptide produced by the
CC salivary gland of the New World sand fly, and streptokinase cDNA

XX Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;

Query Match 54.6%; Score 1145.4; DB 6; Length 8893;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGGCTGTGACCGCTCATCTGTCAACAACAGCAATTGGTGT 647
DB 5935 ATTGCTGACCTGAGTGGCTGTGACCGCTCATCTGTCAACAACAGCAATTGGTGT 5994
QY 648 AGCGTTGCTGGTACCTGTTGAGGGGACGATCAAGACATTAGCTTAAATTTTGAATC 707
DB 5995 AGCGTTGCTGGTACCTGTTGAGGGGACGATCAAGACATTAGCTTAAATTTTGAATC 6054
QY 708 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
DB 6055 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 6114
QY 768 AAACCATTTGCTACTGATAGTGGCGCGATGTCATATAACTTGAGAAAGCTGACTTA 827
DB 6115 AAACCATTTGCTACTGATAGTGGCGCGATGTCATATAACTTGAGAAAGCTGACTTA 6174
QY 828 AAGGCTATTCAAGAACATTGATCGCTAACGTCACAGTAAACGACGTAATTGAGGTC 887
DB 6175 AAGGCTATTCAAGAACATTGATCGCTAACGTCACAGTAAACGACGTAATTGAGGTC 6234
QY 888 ATTGATTTTGCAGACCGATGACACCATTAATCTGATGAAACCGCAAGGCTTACTTGTGAC 947
DB 6235 ATTGATTTTGCAGACCGATGACACCATTAATCTGATGAAACCGCAAGGCTTACTTGTGAC 6294
QY 948 AAAGATGTTGGGTAACTTGGCGAACCCGACCTGTCACAGATTTTGTGAAGGGGACAT 1007
DB 6295 AAAGATGTTGGGTAACTTGGCGAACCCGACCTGTCACAGATTTTGTGAAGGGGACAT 6354
QY 1008 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGCAATCTGTGATGTG 1067
DB 6355 GTGCGGTTAGACCATATAAAGAAAAACAATCAAAACCAAGCAATCTGTGATGTG 6414
QY 1068 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAACAGAGGCTCAAA 1127
DB 6415 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAACAGAGGCTCAAA 6474
QY 1128 GATCTAAGCTATTGAAAACTAGCTATGAGTGAACCATCATCTCAAGAAATTATCTA 1187
DB 6475 GATCTAAGCTATTGAAAACTAGCTATGAGTGAACCATCATCTCAAGAAATTATCTA 6534
QY 1188 GCTCAAGACCAAGCAATTTTAAACAAAACCAACCGCTATACGATTTTGAAGCTGAC 1247
DB 6535 GCTCAAGACCAAGCAATTTTAAACAAAACCAACCGCTATACGATTTTGAAGCTGAC 6594
QY 1248 TCCCTCATGTGTCATGATGACCAATGACATTTTCCGTCGATTTTCAATGATCAAGAG 1307
DB 6595 TCCCTCATGTGTCATGATGACCAATGACATTTTCCGTCGATTTTCAATGATCAAGAG 6654
QY 1308 TTCTACTACCGTGTAAATAATCGGGAACAAGCTTATGAGTCAATAAATAATCTGCTG 1367
DB 6655 TTCTACTACCGTGTAAATAATCGGGAACAAGCTTATGAGTCAATAAATAATCTGCTG 6714
QY 1368 AATGAAGAAATAAACAACCTGATCTCTGAGAAATATTAGTCTTTAAAAAGGG 1427
DB 6715 AATGAAGAAATAAACAACCTGATCTCTGAGAAATATTAGTCTTTAAAAAGGG 6774
QY 1428 GAAAGCCGTATGATCTCTTGAATGCGAGTCACTTGAACCTGTTCAACATCAATACGTT 1487
DB 6775 GAAAGCCGTATGATCTCTTGAATGCGAGTCACTTGAACCTGTTCAACATCAATACGTT 6834
QY 1488 GATGTGATACCAACGATTTGCTAAAAAGTAGAGAGCTTTTAACAGCTAGCGAAGCTAAC 1547

DB 6835 GATGTGATACCAACGATTTGCTAAAAAGTAGAGAGCTTTTAACAGCTAGCGAAGCTAAC 6894
QY 1548 TTAACTTCAGAGATTTTATACGATCTCTGATATAGGCTTAAGCTTAACCTCTTACAAATCTC 1607
DB 6895 TTAACTTCAGAGATTTTATACGATCTCTGATATAGGCTTAAGCTTAACCTCTTACAAATCTC 6954
QY 1608 GATGCTTTTGGTATTTATAGACTATACCTTAACCTGAAAAAGTAGAGATTAATCAGATGAC 1667
DB 6955 GATGCTTTTGGTATTTATAGACTATACCTTAACCTGAAAAAGTAGAGATTAATCAGATGAC 7014
QY 1668 ACCAAGCTATCATACCGTTTATATGAGCAAGGACCCGAAAGAGAGAAATGCTAGCTAC 1727
DB 7015 ACCAAGCTATCATACCGTTTATATGAGCAAGGACCCGAAAGAGAGAAATGCTAGCTAC 7074
QY 1728 CATTTAGCTGTGCT 1742
DB 7075 CATTTAGCTGTGAT 7089

RESULT 11

AAK1632 standard; DNA; 1242 BP.

XX AAK1632;
AC 17-OCT-2003 (revised)
DT 04-MAY-1999 (first entry)
XX Streptococcus equisimilis native streptokinase encoding DNA.

KW Streptococcus equisimilis H46a; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance; de.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers

FT 1. 1242 /tag= a

FT /transl_except= (pos:40..42,aa:Aen)

FT /note= "no stop codon given"

XX US5876999-A.

XX 02-MAR-1999.

XX 06-DEC-1995; 95US-00568393.

XX 06-DEC-1995; 95US-00568393.

XX (NASC-) NAT SCI COUNCIL.

XX Wu H;

XX WPI; 1999-189643/16.

XX P-PSDB; AAW94664.

XX Mutant streptokinase polypeptide - useful as plasmin-resistant

XX thrombolytic agent.

XX Claim 1; Col 7-10; 17p; English.

CC The present invention describes a mutant streptokinase (SK) polypeptide
CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
CC segment of the corresponding native SK is replaced by another amino acid.
CC The present sequence encodes native SK. SK is a secretory protein of
CC haemolytic Streptococcus able to activate human plasminogen (HPIg) to
CC plasmin (HPIa), which is a serine protease able to catalyse the
CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
CC agent in the treatment of vascular thromboembolytic symptoms such as

Best Local Similarity 99.4%; Pred. No. 0;
Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 588 ATTCGCGAGCCTGAGCGCTGCTAGACCCCTGCATCTGTCAACAACACCGCAATGTTGTT 647
Db 358 ATTCGCGAGCCTGAGCGCTGCTAGACCCCTGCATCTGTCAACAACACCGCAATGTTGTT 417
QY 648 AGCGTTGCTGCTGCTGCTGAGGGGACGATCAAGACATTAAGTCTTAATTTTGAATTC 707
Db 418 AGCGTTGCTGCTGCTGCTGAGGGGACGATCAAGACATTAAGTCTTAATTTTGAATTC 477
QY 708 GATCTAACATCAGACCTGCTCATGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 767
Db 478 GATCTAACATCAGACCTGCTCATGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 537
QY 768 AAACCAATTTGCTACTGATAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTTACTA 827
Db 538 AAACCAATTTGCTACTGATAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTTACTA 597
QY 828 AAGGCTATTCAAGAACAAATGATGCTAAGTCCAGAGTAAAGCGACTAATTGAGAGTC 887
Db 598 AAGGCTATTCAAGAACAAATGATGCTAAGTCCAGAGTAAAGCGACTAATTGAGAGTC 657
QY 888 ATTCGATTTTTCGAGCGATGCAACCATTAATGATCGAAACGCGCAAGTCTTAATTGCTGAC 947
Db 658 ATTCGATTTTTCGAGCGATGCAACCATTAATGATCGAAACGCGCAAGTCTTAATTGCTGAC 717
QY 948 AAAGATGCTTCGGTAACTTTCCGAGCCCAACTGCTCCAGAAATTTTGTCTAAGCGGACAT 1007
Db 718 AAAGATGCTTCGGTAACTTTCCGAGCCCAACTGCTCCAGAAATTTTGTCTAAGCGGACAT 777
QY 1008 GTGCGGGTTAGACCATTAAGAAAAACAATCAAAACCAAGGAAATCTGTTGAGTGTG 1067
Db 778 GTGCGGGTTAGACCATTAAGAAAAACAATCAAAACCAAGGAAATCTGTTGAGTGTG 837
QY 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGGTCTCAA 1127
Db 838 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGGTCTCAA 897
QY 1128 GATTAAGTAACTTAAAGAAACATAGCTATGCTGTAACCATCAATCATCTCAAGAAATTA 1187
Db 898 GATTAAGTAACTTAAAGAAACATAGCTATGCTGTAACCATCAATCATCTCAAGAAATTA 957
QY 1188 GCTGAAGCAAGCAATTTTAAACAAACCAACCCGCGCTATGATTTTGAACCGTGAC 1247
Db 958 GCTGAAGCAAGCAATTTTAAACAAACCAACCCGCGCTATGATTTTGAACCGTGAC 1017
QY 1248 TCCCTCAATCGTCACTGATGACATGACATTTTCCGTAAGATTTTCAAGATCAAGAG 1307
Db 1018 TCCCTCAATCGTCACTGATGACATGACATTTTCCGTAAGATTTTCAAGATCAAGAG 1077
QY 1308 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAATCTGGTCTG 1367
Db 1078 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAAATCTGGTCTG 1137
QY 1368 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGGG 1427
Db 1138 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGGG 1197
QY 1428 GAAAAACCCGTATATCCCTTGTATGCGAGTCACTTGAACCTGTTCAACATCAATAAGTT 1487
Db 1198 GAAAAACCCGTATATCCCTTGTATGCGAGTCACTTGAACCTGTTCAACATCAATAAGTT 1257
QY 1488 GATGTGATACCAAGAAATGCTAAATAAGTGAAGCTCTTAAACGCTAGCGAAGCTTAC 1547
Db 1258 GATGTGATACCAAGAAATGCTAAATAAGTGAAGCTCTTAAACGCTAGCGAAGCTTAC 1317
QY 1548 TTAGACTTCAGAGATTTATTCGATCTCTGATTAAGGCTTAACTACTCTCAACAATCTC 1607
Db 1318 TTAGACTTCAGAGATTTATTCGATCTCTGATTAAGGCTTAACTACTCTCAACAATCTC 1377
QY 1608 GATGCTTTGGTATTTATGAGCTATACCTTAACGGAAGGAGTAATCAAGATGAC 1667

```

```

Db 1378 GATGCTTTGGTATTTATGAGCTATACCTTAACGGAAGGAGATTAATCAAGATGAC 1437
QY 1668 ACCAACCGTATCATTAACCGTTTATATGGCAAGGACCCGGAAGAGAGAAATGCTAGTAC 1727
Db 1438 ACCAACCGTATCATTAACCGTTTATATGGCAAGGACCCGGAAGAGAGAAATGCTAGTAT 1497
QY 1728 CATTTAGCTGGTGGT 1742
Db 1498 CATTTAGCTATGAT 1512

```

RESULT 13

AAAX16633
ID AAAX16633 standard; DNA; 1242 BP.

XX AAX16633;

DT 04-MAY-1999 (first entry)

XX Streptococcus equisimilis mutant streptokinase K59E encoding DNA.

XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;

XX plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;

XX serine protease; fibrin; blood clot; thrombolytic;

XX vascular thromboembolytic symptom; acute myocardial infarction;

XX fibrinolysis; resistance; ds.

OS Streptococcus dysgalactiae subsp. equisimilis.

XX Synthetic.

XX Key

PD 02-MAR-1999.

XX 06-DEC-1995; 95US-00568393.

XX 06-DEC-1995; 95US-00568393.

XX (NASC-) NAT SCI CODNCLL.

XX Wu H;

XX WPI; 1999-189643/16.

DR P-PSDB; AAX94665.

XX Mutant streptokinase polypeptide - useful as plasmin-resistant

PT thrombolytic agent.

XX Claim 4; Col 11-14; 17pp; English.

XX The present invention describes a mutant streptokinase (SK) polypeptide

CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61

CC segment of the corresponding native SK is replaced by another amino acid.

CC The present sequence encodes mutant SK K59E. SK is a secretory protein of

CC haemolytic Streptococcus able to activate human plasminogen (Hpi) to

CC plasmin (HpiM), which is a serine protease able to catalyse the

CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic

CC agent in the treatment of vascular thromboembolytic symptoms such as

CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant

CC both in acting as a fibrinolytic agent and in activating human plasminogen

XX Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;

Query Match 54.4%; Score 1140.6; DB 2; Length 1242;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1146; Conservative 0; Mismatches 9; Indels 0; Gaps 0;


```

QY 588 ATTGCTGACCTGAGTGGCTGCTGAGACCGTCCATCTGTCAACAAGCCCAATGGTGT 647
DB 1 ATTGCTGACCTGAGTGGCTGCTGAGACCGTCCATCTGTCAACAAGCCCAATGGTGT 60
QY 648 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGCATTAAGCTTAAATTTTGAATC 707
DB 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGCATTAAGCTTAAATTTTGAATC 120
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
DB 121 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCGAATCA 180
QY 768 AAACCATTTGCTAGTATGAGGCGCATGTCATTAACCTTGAAGAAAGCTGACTTACTA 827
DB 181 AAACCATTTGCTAGTATGAGGCGCATGTCATTAACCTTGAAGAAAGCTGACTTACTA 240
QY 828 AAGGCTTATCAAGAAACAATTGATGCTTACCTCCAGTAAAGAGCTACTTGAAGTC 887
DB 241 AAGGCTTATCAAGAAACAATTGATGCTTACCTCCAGTAAAGAGCTACTTGAAGTC 300
QY 888 ATTGATTTTGAAGGATGCAACCATTAAGTGAAGGGAAGGCTTACTTGGCTGAC 947
DB 301 ATTGATTTTGAAGGATGCAACCATTAAGTGAAGGGAAGGCTTACTTGGCTGAC 360
QY 948 AAAGATGTTTGGTAACTTGGCCGACCCGACCTGTCCAGAAATTTTGTCTAAGCGGACAT 1007
DB 361 AAAGATGTTTGGTAACTTGGCCGACCCGACCTGTCCAGAAATTTTGTCTAAGCGGACAT 420
QY 1008 GTGCGGCTTGAACCATTAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 1067
DB 421 GTGCGGCTTGAACCATTAAGAAACCAATACAAACCAAGGAAATCTGTTGATGTG 480
QY 1068 GAATATATCTGATGATTAACCTCTTAAACCTGATGAGATTTTCAAGCCAGTCTCAA 1127
DB 481 GAATATATCTGATGATTAACCTCTTAAACCTGATGAGATTTTCAAGCCAGTCTCAA 540
QY 1128 GATATCTAAGCTATTTGAAACCACTAGCTATCGGTGACACCATCACTCAAGAAATTA 1187
DB 541 GATATCTAAGCTATTTGAAACCACTAGCTATCGGTGACACCATCACTCAAGAAATTA 600
QY 1188 GCTGAAGCAAAAGCATTTTAAACAAACCAAGGCTATAGCATTTTAAAGCTGAC 1247
DB 601 GCTGAAGCAAAAGCATTTTAAACAAACCAAGGCTATAGCATTTTAAAGCTGAC 660
QY 1248 TCTCTAATGCTCACTGATGACATGACATTTCCGTAGCATTTTAAACATGATCAAG 1307
DB 661 TCTCTAATGCTCACTGATGACATGACATTTCCGTAGCATTTTAAACATGATCAAG 720
QY 1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGCATTTTAAAAATCTGCTG 1367
DB 721 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGCATTTTAAAAATCTGCTG 780
QY 1368 AATGAAGAAATAAACAACACTGACCTGATCTGAGAAATTTTACGTCCTTAAAAAGG 1427
DB 781 AATGAAGAAATAAACAACACTGACCTGATCTGAGAAATTTTACGTCCTTAAAAAGG 840
QY 1428 GAAAAGCCGTATGATCCCTTTGATGCAAGTCACTTGAAGCTTCAACCAATACGTT 1487
DB 841 GAAAAGCCGTATGATCCCTTTGATGCAAGTCACTTGAAGCTTCAACCAATACGTT 900
QY 1488 GATGTGATACCAACGAATGTCTTAAAGGAGAGAGCTTTTAAAGCTAGGAAAGCTAAC 1547
DB 901 GATGTGATACCAACGAATGTCTTAAAGGAGAGAGCTTTTAAAGCTAGGAAAGCTAAC 960
QY 1548 TTAGACTTCAGAGATTTTACGATCTCTGATTAAGGCTTAACTACTTCAACCAATCTC 1607
DB 961 TTAGACTTCAGAGATTTTACGATCTCTGATTAAGGCTTAACTACTTCAACCAATCTC 1020
QY 1608 GATGCTTTTGGTATTTATGACATTAACCTTAACTGAAAGAGTGAAGATTAACGATGAC 1667
DB 1021 GATGCTTTTGGTATTTATGACATTAACCTTAACTGAAAGAGTGAAGATTAACGATGAC 1080

```

```

QY 1668 ACCAACCGTATCATTAACGTTTATATGAGGCAAGCCGAGAGAGATGCTACTAC 1727
DB 1081 ACCAACCGTATCATTAACGTTTATATGAGGCAAGCCGAGAGAGATGCTACTAC 1140
QY 1728 CATTTAGCTGTGTGT 1742
DB 1141 CATTTAGCTGTGTAT 1155

```

```

RESULT 14
AAA37628
ID AAA37628 standard; DNA; 1327 BP.
XX
XX AAA37628;
AC 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
DT
XX Streptokinase-NTR gene.
DE
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
XX Streptococcus dysgalactiae subsp. equisimilis.
OS
XX BP1024192-A2.
PN
XX 02-AUG-2000.
PD
XX 23-DEC-1999; 99EP-00310541.
PF
XX 24-DEC-1998; 98IN-DE003825.
PR
XX (COUL ) CSIR COUNCIL SCI IND RES.
PA
XX
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
PI WPI; 2000-516032/47.
DR
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
XX Example 1; Fig 14; 58bp; English.
PS
XX
XX This sequence represents a streptokinase-NTR (SK-NTR) gene (where NTR
CC stands for N-terminally repaired with native sequence). The invention
CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
CC fusion between streptokinase (SK), which are capable of plasminogen (Pg)
CC activation, and fibrin binding regions of human fibrinectin, which are
CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
CC possesses the ability to bind with fibrin independently and also
CC characteristically retains a Pg activation ability which becomes evident
CC only after a pronounced duration, or lag, after exposure of the PA to a
CC suitable animal or human Pg. The hybrid streptokinase-fibrin binding
CC domain polypeptides are useful in thrombolytic therapy for various kinds
CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
CC as well as kinetics of plasminogen activation that are distinct from that
CC of natural streptokinase in being characterised by a temporary delay, or
CC lag of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolytics). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
XX Sequence 1327 BP; 456 A; 279 C; 250 G; 342 T; 0 U; 0 Other;
SQ

```

Query Match 54.1%; Score 1134.2; DB 3; Length 1327;

Best Local Similarity 98.9%; Pred. No. 2.3e-313;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

QY 588 ATTGCTGACCTGATGCGCTGCTAGACCTGCTCACTCTGCAACACAGCGCAATGGTGT 647
Db 83 ATAGCTGCTCTTAATGCTACTAGATCGCTCTTCTTAATAACAGCGCAATGGTGT 142
QY 648 AGGTTGCTGCTGCTAGTGAAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 707
Db 143 AGGTTGCTGCTGCTAGTGAAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 202
QY 708 GATCTAACATCAGACCTGCTCATGAGAAAGACAGACAGGCTTAAAGTCAAAAATCA 767
Db 203 GATCTAACATCAGACCTGCTCATGAGAAAGACAGACAGGCTTAAAGTCAAAAATCA 262
QY 768 AAACCACTTCTCTACTGATAGTGGCGCATGTCACTATACTTGAGAAAGTGACTTACTA 827
Db 263 AAACCACTTCTCTACTGATAGTGGCGCATGTCACTATACTTGAGAAAGTGACTTACTA 322
QY 828 AAGGCTATTCAAGAACAAATTGATGCTAAGTCCAGCTAAGCGACTTGAAGTGC 887
Db 323 AAGGCTATTCAAGAACAAATTGATGCTAAGTCCAGCTAAGCGACTTGAAGTGC 382
QY 888 ATTGATTTTGAAGCGATGCAACCATTAATGATGAAACGGCAAGTCTACTTTGCTGAC 947
Db 383 ATTGATTTTGAAGCGATGCAACCATTAATGATGAAACGGCAAGTCTACTTTGCTGAC 442
QY 948 AAAGATGCTTCCGTAACCTTCCGACCCAACTCTGTCCAAGAAATTTTGTAAAGCGACAT 1007
Db 443 AAAGATGCTTCCGTAACCTTCCGACCCAACTCTGTCCAAGAAATTTTGTAAAGCGACAT 502
QY 1008 GTGCGGGTTAGACATATTAAGAAAAACAATACAAAACCAAGGAAATCTGTGATGTG 1067
Db 503 GTGCGGGTTAGACATATTAAGAAAAACAATACAAAACCAAGGAAATCTGTGATGTG 562
QY 1068 GAATATCTGTATCAGTTTACTCTCTTAAACCCGTATGACGATTTCAAGACGAGTCTCAA 1127
Db 563 GAATATCTGTATCAGTTTACTCTCTTAAACCCGTATGACGATTTCAAGACGAGTCTCAA 622
QY 1128 GATACTAAGCTATTTGAAAAACACTAGCTATCGGTGACACATCATCTCAAGAAATTA 1187
Db 623 GATACTAAGCTATTTGAAAAACACTAGCTATCGGTGACACATCATCTCAAGAAATTA 682
QY 1188 GCTCAAGCACAAAGCAATTTTAAACAAAACCAACCGCTATATGATTTTGAACGTGAC 1247
Db 683 GCTCAAGCACAAAGCAATTTTAAACAAAACCAACCGCTATATGATTTTGAACGTGAC 742
QY 1248 TCCCTCAATGCTCATCATGACAAATGACATTTTCCGTACGATTTTAAACCAATGATCA 1307
Db 743 TCCCTCAATGCTCATCATGACAAATGACATTTTCCGTACGATTTTAAACCAATGATCA 802
QY 1308 TTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGATCATATAAAAAATCGTCTG 1367
Db 803 TTTACTTACCGTGTAAAAATCGGGAAACAAGCTTATAGATCATATAAAAAATCGTCTG 862
QY 1368 AATGAAGAATTAACAACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAAGGG 1427
Db 863 AATGAAGAATTAACAACTGACCTGATCTCTGAGAAATATTAAGTCTTAAAAAAGGG 922
QY 1428 GAAAAACCCGTATATCCCTTGTATGCGAGTCACTTGAACGTTTCAACATCAAAATAGGTT 1487
Db 923 GAAAAACCCGTATATCCCTTGTATGCGAGTCACTTGAACGTTTCAACATCAAAATAGGTT 982
QY 1488 GATGTCGATACCAAGCAATTTGCTAAAAAGTGAAGCACTTTAACAGCTAGGAAAGTAA 1547
Db 983 GATGTCGATACCAAGCAATTTGCTAAAAAGTGAAGCACTTTAACAGCTAGGAAAGTAA 1042
QY 1548 TTAGACTTCAGAGATTTATAGATCTCTGATATAGGCTTAACTACTCTCAACAATCTC 1607
Db 1043 TTAGACTTCAGAGATTTATAGATCTCTGATATAGGCTTAACTACTCTCAACAATCTC 1102
QY 1608 GATGCTTTGTATATAGATCTATACCTTAACTGAAAAGTAAAGGATTAATCAGATGAC 1667

```

```

Db 1103 GATGCTTTGGTATATGACCTATACCTTAACTGGAAGAAAGTAAAGATATCAGATGAC 1162
QY 1668 ACCAACCTTATCATACCGTTTATATGGGCAAGCAACCCGAAGAGAAATGCTAGTAC 1727
Db 1163 ACCAACCTTATCATACCGTTTATATGGGCAAGCAACCCGAAGAGAAATGCTAGTAT 1222
QY 1728 CATTTAGCTGGTGT 1742
Db 1223 CATTTAGCTATGAT 1237

```

RESULT 15

AAT77778
ID AAT77778 standard; cDNA; 2566 BP.

AA777778;

17-OCT-2003 (revised)

01-OCT-1997 (first entry)

Coding sequence for plasminogen-binding fragment of Streptokinase.

Plasminogen-binding fragment; streptokinase; degradation; MBP; thrombolytic agent; blood clot; bolus; maltose-binding protein; ds.

Streptococcus dysgalactiae subsp. equisimilis.

W09641883-A1.

27-DEC-1996.

07-JUN-1996; 96WO-US009640.

09-JUN-1995; 95US-00488940.

(HARD) HARVARD COLLEGE.

Reed GL;

WPI; 1997-065469/06.

Modified forms of streptokinase resistant to enzymatic cleavage - useful as thrombolytic agents in treating thrombosis and in medical equipment.

Example 1; Page 22-23; 65pp; English.

This sequence encodes the wild type plasminogen-binding fragment of streptokinase. The protein fragment encoded by this sequence was used in the design of a modified streptokinase has an in vitro degradation rate at least 2 times slower than that of native streptokinase. Compounds containing modified streptokinases are specifically used as thrombolytic agents for dissolving blood clots in vivo in a mammal, preferably at a dose of 20000 U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to standardise OS field)

Sequence 2566 BP; 825 A; 545 C; 469 G; 727 T; 0 U; 0 Other;

Query Match 54.0%; Score 1132.4; DB 2; Length 2566;

Best Local Similarity 99.3%; Pred. No. 1.1e-312;
Matches 1148; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

```

QY 587 TATTGCTGACCTGATGAGCTGCTAGACCGTCACTCTGTCAACACAGCCAAATGGTGT 646
Db 896 TATTGCTGACCTGATGAGCTGCTAGACCGTCACTCTGTCAACACAGCCAAATGAGTGT 955
QY 647 TAGGTTGCTGCTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 706
Db 956 TAGGTTGCTGCTACTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAAT 1015
QY 707 CGATCTAACATCAGACCTGCTCATGAGAAAGACAGAGCAAGGCTTAAAGTCAAAATC 766
Db 1016 CGATCTAACATCAGACCTGCTCAT--AGAAAGACAGAGCAAGGCTTAAAGTCAAAATC 1073

```

QY 767 AAAACCACTTGTCTATCTGATAGTGGCGGAGTGTCAATTAACCTTGAGAAAGCTGACTTACT 826
| | | | |
Db 1074 AAAACCACTTGTCTATCTGATAGTGGCGGAGTGTCAATTAACCTTGAGAAAGCTGACTTACT 1133
| | | | |
QY 827 AAAGGCTATTCAGAAACAATTGATCGTAAAGTCCACGTAAACGACGACTACTTTGAGGT 886
| | | | |
Db 1134 AAAGGCTATTCAGAAACAATTGATCGTAAAGTCCACGTAAACGACGACTACTTTGAGGT 1193
| | | | |
QY 887 CATGTATTTGCAAGCGATGCAACCATTTAATGATCGAAACGCAAGGTCTACTTTGCTGA 946
| | | | |
Db 1194 CATGTATTTGCAAGCGATGCAACCATTTAATGATCGAAACGCAAGGTCTACTTTGCTGA 1253
| | | | |
QY 947 CAAAGATGTTCCGTTAACTTTCGCGACCCAACTGTCCAGAAATTTTGTAAAGCGACA 1006
| | | | |
Db 1254 CAAAGATGTTCCGTTAACTTTCGCGACCCAACTGTCCAGAAATTTTGTAAAGCGACA 1313
| | | | |
QY 1007 TGTGCGCGTTAGACCATATTAAGAAACCAATACAAACCAAGGAAATCTGTTGATGT 1066
| | | | |
Db 1314 TGTGCGCGTTAGACCATATTAAGAAACCAATACAAACCAAGGAAATCTGTTGATGT 1373
| | | | |
QY 1067 GGAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGACGACGAGTCTCAA 1126
| | | | |
Db 1374 GGAATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGACGACGAGTCTCAA 1433
| | | | |
QY 1127 AGATTAAGCTATTTGAAAACTAGCTATCGTGAACAATCAATCTCAAGAAATTACT 1186
| | | | |
Db 1434 AGATTAAGCTATTTGAAAACTAGCTATCGTGAACAATCAATCTCAAGAAATTACT 1493
| | | | |
QY 1187 AGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTATGAACGTGA 1246
| | | | |
Db 1494 AGCTCAAGCAAAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTATGAACGTGA 1553
| | | | |
QY 1247 CTGCTCAATGCTGACATGACATTTGCGTACGATTTTACCAATGATCAAGA 1306
| | | | |
Db 1554 CTGCTCAATGCTGACATGACATTTGCGTACGATTTTACCAATGATCAAGA 1613
| | | | |
QY 1307 GTTTACTTACCGTGTAAAAATCGGAAACAAGTTATAGATCAATAAAAAATCTGCTCT 1366
| | | | |
Db 1614 GTTTACTTACCGTGTAAAAATCGGAAACAAGTTATAGATCAATAAAAAATCTGCTCT 1673
| | | | |
QY 1367 GAAATGAAGAAATTAACAACAATGACTGATCTGAGAAATATTACGCTTTAAAAAAG 1426
| | | | |
Db 1674 GAAATGAAGAAATTAACAACAATGACTGATCTGAGAAATATTACGCTTTAAAAAAG 1733
| | | | |
QY 1427 GGAAGAGCGGTATGATCCCTTTGATGCGAGTCACTTGAAACTGTTCAACCAATACGT 1486
| | | | |
Db 1734 GGAAGAGCGGTATGATCCCTTTGATGCGAGTCACTTGAAACTGTTCAACCAATACGT 1793
| | | | |
QY 1487 TGAATGTGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTTTAAAGCTAGCGAACGTA 1546
| | | | |
Db 1794 TGAATGTGATACCAACGAATTTGCTAAAAAGTGAAGCAGCTTTAAAGCTAGCGAACGTA 1853
| | | | |
QY 1547 CTTTGAATTCAGAGATTTTATGATCCTCGTATAGGCTTAACTTACTTCAACAATCT 1606
| | | | |
Db 1854 CTTTGAATTCAGAGATTTTATGATCCTCGTATAGGCTTAACTTACTTCAACAATCT 1913
| | | | |
QY 1607 CGATGCTTTTGGTATTAATGACATATACCTTTAATCTGAAAAAGTGAAGATTAATCAAGATGA 1666
| | | | |
Db 1914 CGATGCTTTTGGTATTAATGACATATACCTTTAATCTGAAAAAGTGAAGATTAATCAAGATGA 1973
| | | | |
QY 1667 CACCAACCGGTATCATTAACGTTTATATGCGCAAGCGAACCCGAAAGAGAAATGCTAGCTA 1726
| | | | |
Db 1974 CACCAACCGGTATCATTAACGTTTATATGCGCAAGCGAACCCGAAAGAGAAATGCTAGCTA 2033
| | | | |
QY 1727 CCAATTAAGCTGTGTGT 1742
| | | | |
Db 2034 TCAATTTAGCCTATGAT 2049
| | | | |

QY	768	AAACATTTCGACATGATAGTGGCCGATGTCACATTAACCTTGAGAAAGCTGACTTACGA	827
Db	181	AAACCAATTTCGACATGATAGTGGCCGATGTCACATTAACCTTGAGAAAGCTGACTTACGA	240
QY	828	AAGGCTATTCAAGAACAAATGATGCTTAACGTCACAAGTAACGACGACTACTTGAGGTC	887
Db	241	AAGGCTATTCAAGAACAAATGATGCTTAACGTCACAAGTAACGACGACTACTTGAGGTC	300
QY	888	ATTGATTTTTCGAAGGATGCAACCACTTAACCTGATTCGAAAACGGCAAGGCTACTCTTGGTGA	947
Db	301	ATTGATTTTTCGAAGGATGCAACCACTTAACCTGATTCGAAAACGGCAAGGCTACTCTTGGTGA	360
QY	948	AAAGATGCTCGGTACCTTGGCCGACCCCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT	1007
Db	361	AAAGATGCTCGGTACCTTGGCCGACCCCAACCTGTCCAGAAATTTTTCGTAAGCGGACAT	420
QY	1008	GTGCGCGCTTGAACCACTTAATGAAGAAAACCAATTAACAACCAAGCAAAATCTGTTGATGTG	1067
Db	421	GTGCGCGCTTGAACCACTTAATGAAGAAAACCAATTAACAACCAAGCAAAATCTGTTGATGTG	480
QY	1068	GAATATACCTGATCAGTTTACTCCCTTAACCCGTATGACGATTTAGACCAAGGCTTCAAA	1127
Db	481	GAATATACCTGATCAGTTTACTCCCTTAACCCGTATGACGATTTAGACCAAGGCTTCAAA	540
QY	1128	GATATCTAAGCTATTGAAAACACTAGCTATCGGTGACCAACATCAGATCTCAAGAAATTACTA	1187
Db	541	GATATCTAAGCTATTGAAAACACTAGCTATCGGTGACCAACATCAGATCTCAAGAAATTACTA	600
QY	1188	GCTCAGACACAAAGATTTTAAACAACCAACCAAGGCTATGCAATTTATGAAGGTGAC	1247
Db	601	GCTCAGACACAAAGATTTTAAACAACCAACCAAGGCTATGCAATTTATGAAGGTGAC	660
QY	1248	TCCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGACATTTTTCGAATGATCAAGAG	1307
Db	661	TCCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAGACATTTTTCGAATGATCAAGAG	720
QY	1308	TTTACTTAACCGTGTTAAAAATCGGGAAACAAGCTTATAGATCAATTAATAAAATCTGGCTG	1367
Db	721	TTTACTTAACCGTGTTAAAAATCGGGAAACAAGCTTATAGATCAATTAATAAAATCTGGCTG	780
QY	1368	AATATGAAGAAATTAACAACAACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGG	1427
Db	781	AATATGAAGAAATTAACAACAACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGG	840
QY	1428	GAAAAGCCGTATGATCCCTTGGATGCGACGTCACTTGAACCTGTTCAACCATCAAAATACGTT	1487
Db	841	GAAAAGCCGTATGATCCCTTGGATGCGACGTCACTTGAACCTGTTCAACCATCAAAATACGTT	900
QY	1488	GATGTCGATACCAAGCAATTCCTAAAAAGTAGACAGCTCTTAAACGCTAGCGAAGCTAAC	1547
Db	901	GATGTCGATACCAAGCAATTCCTAAAAAGTAGACAGCTCTTAAACGCTAGCGAAGCTAAC	960
QY	1548	TTTAACTTCAAGATTTTATAGATCCCTGCTGATTAAGGCTAAACCTTCAACAACAATTC	1607
Db	961	TTTAACTTCAAGATTTTATAGATCCCTGCTGATTAAGGCTAAACCTTCAACAACAATTC	1020
QY	1608	GATGCTTTGGTATTTATGACTATACCTTAACTGGAAGAAAGTAGAGATATCAAGATGAC	1667
Db	1021	GATGCTTTGGTATTTATGACTATACCTTAACTGGAAGAAAGTAGAGATATCAAGATGAC	1080
QY	1668	ACCAACCGTATCATACCGTTTATATGAGGCAAGCCGACGAAGAGAGAAATGCTAGCTAC	1727
Db	1081	ACCAACCGTATCATACCGTTTATATGAGGCAAGCCGACGAAGAGAGAAATGCTAGCTAT	1140
QY	1728	CATTAGCTAGTGGGT 1742	
Db	1141	CATTAGCTATGAT 1155	

RESULT 3
US-08-568-393B-1
; Sequence 1, Application US/08568393B
; Patent No. 5876999

```

1 GENERAL INFORMATION:
2 APPLICANT: Hua-Lin Wu
3 TITLE OF INVENTION: Preparation of novel streptokinase
4 TITLE OF INVENTION: mutants as improved thrombolytic agents
5 NUMBER OF SEQUENCES: 2
6 CORRESPONDENCE ADDRESS:
7 ADDRESS: Jeling & Chang
8 STREET: Two No. 5876995th Second Street, Suite 290
9 CITY: San Jose
10 STATE: California
11 COUNTRY: USA
12 ZIP: 95113
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
16 MEDIUM TYPE: storage
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Wordperfect 6.1 on Windows 3.1
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/568,393B
22 FILING DATE:
23 CLASSIFICATION: 435
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Chi-Ping Chang
26 REGISTRATION NUMBER: 37,798
27 REFERENCE/DOCKET NUMBER:
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (408) 288-8585
30 TELEFAX: (408)288-8386
31 INFORMATION FOR SEQ ID NO. 1:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 1242 base pairs
34 TYPE: Nucleic Acid
35 STRANDEDNESS: double
36 TOPOLOGY: linear
37
38 MOLECULE TYPE:
39 HYPOTHEICAL: N
40 ANTI-SENSE: N
41 ORIGINAL SOURCE:
42 ORGANISM: Streptococcus equisimilis H46A
43 INDIVIDUAL ISOLATE: "Malke, H., Roe, B., and Ferretti, J. J.,"
44 INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from
45 INDIVIDUAL ISOLATE: equisimilis H46A" from Gene 34:357-362 (1985).
46 CELL TYPE: Streptococcus equisimilis H46A
47
48 US-08-568-393B-1
49
50 Query Match 54.6%; Score 1143.8; DB 2; Length 1242;
51 Best Local Similarity 99.4%; Pred. No. 0;
52 Matches 1148; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```

```

Db 301 ATTGATTTTGCAGCGATGCAACCATTAAGTCAAAACGCGAGCTCTAATTCTGAC 360
QY 948 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 1007
Db 361 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 420
QY 1008 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAACCAAGGAAATCTGTGATGTG 1067
Db 421 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAACCAAGGAAATCTGTGATGTG 480
QY 1068 GAATATATCTGTACATTTACTCTCCCTTAACCCCTGATGACATTTTCAAGCCAGCTCAAA 1127
Db 481 GAATATATCTGTACATTTACTCTCCCTTAACCCCTGATGACATTTTCAAGCCAGCTCAAA 540
QY 1128 GATCTAAGCTATTTGAAAACATAGCATATGCTGACCACTCAATCTCAAGAAATTTCTA 1187
Db 541 GATCTAAGCTATTTGAAAACATAGCATATGCTGACCACTCAATCTCAAGAAATTTCTA 600
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGCTATACGATTTTAAAGCTGAC 1247
Db 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCCAGCTATACGATTTTAAAGCTGAC 660
QY 1248 TCCTCAATGCTCACTGATGCAATGACATTTTCCGATCGATTTTACCAATGATCAAAG 1307
Db 661 TCCTCAATGCTCACTGATGCAATGACATTTTCCGATCGATTTTACCAATGATCAAAG 720
QY 1308 TTTACTTACCGGTATTAATAATCCGGAAACAGCTTATAGATCAATTAATAATCTGTCTG 1367
Db 721 TTTACTTACCGGTATTAATAATCCGGAAACAGCTTATAGATCAATTAATAATCTGTCTG 780
QY 1368 AATGAAGAAATTAACAACATGACCTGATCTGAGAAATTTACGTCTTAAAAAAGG 1427
Db 781 AATGAAGAAATTAACAACATGACCTGATCTGAGAAATTTACGTCTTAAAAAAGG 840
QY 1428 GAAAGCCGTATGATCCCTTTGATGCGACATCTTGAACCTGTTCAACATCAAAATGCTT 1487
Db 841 GAAAGCCGTATGATCCCTTTGATGCGACATCTTGAACCTGTTCAACATCAAAATGCTT 900
QY 1488 GATGTGATTCACACGATTTGCTTAAAGAGGAGGAGCTCTTAACAGCTAGGAAAGTAC 1547
Db 901 GATGTGATTCACACGATTTGCTTAAAGAGGAGGAGGAGCTCTTAACAGCTAGGAAAGTAC 960
QY 1548 TTAGACTTCAGAGATTTATACGATCCCTGATGTAAGGCTAACTACTCTAACAAATCTC 1607
Db 961 TTAGACTTCAGAGATTTATACGATCCCTGATGTAAGGCTAACTACTCTAACAAATCTC 1020
QY 1608 GATGCTTTGGTATTTATGACTATACCTTAACCTGAAAAGTAGAGATTAATCAGATGAC 1667
Db 1021 GATGCTTTGGTATTTATGACTATACCTTAACCTGAAAAGTAGAGATTAATCAGATGAC 1080
QY 1668 ACCAACCCTATCAATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGTAC 1727
Db 1081 ACCAACCCTATCAATACCGTTTATATGCGCAAGCGACCCGAGAGAGAGATGCTAGTAC 1140
QY 1728 CATTTAGCTGTGTGT 1742
Db 1141 CATTTAGCTGTGTGT 1155

```

RESULT 4
US-08-568-393B-2

Sequence 2, Application US/08568393B
Patent No. 5876999
GENERAL INFORMATION:
APPLICANT: Hua-Lin Wu
TITLE OF INVENTION: Preparation of novel streptokinase
TITLE OF INVENTION: mutants as improved thrombolytic agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jeing & Chang
STREET: Two No. 5876999th Second Street, Suite 290
CITY: San Jose

```

STATE: California
COUNTRY: USA
ZIP: 95113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
MEDIUM TYPE: Storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1 on Window 3.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,393B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chi-Ping Chang
REGISTRATION NUMBER: 37,798
TELEPHONE: (408) 288-8585
TELEFAX: (408) 288-8386
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1242 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: N
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: SK-K59E
LOCATION: DNA sequence No. 5876999174 and 175 have been changed
LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
LOCATION: from Lys to Glu.
OTHER INFORMATION:
US-08-568-393B-2

Query Match 54.4% Score 1140.6; DB 2; Length 1242;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1146; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGTGCTGACCGCTCATCTGTCAACACAGCAATGTGTTT 647
Db 1 ATTGCTGACCTGAGTGTGCTGACCGCTCATCTGTCAACACAGCAATGTGTTT 60
QY 648 AGCGTGTGCTGATCTGTTGAGGGGACGATCAAGCATTAATTTTGAATTC 707
Db 61 AGCGTGTGCTGATCTGTTGAGGGGACGATCAAGCATTAATTTTGAATTC 120
QY 708 GATCTAATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 767
Db 121 GATCTAATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 180
QY 121 GATCTAATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 180
Db 768 AAACATTTGCTACTGATAGTGGCGGATGTCACATTAACCTGAGAAAGCTGACTTACTA 827
QY 181 AAACATTTGCTACTGATAGTGGCGGATGTCACATTAACCTGAGAAAGCTGACTTACTA 240
Db 828 AAGCTATTCAGAAACAAATTAATGCTTAACGTCACAGTAAAGCACTACTTTGAGGTC 887
QY 241 AAGCTATTCAGAAACAAATTAATGCTTAACGTCACAGTAAAGCACTACTTTGAGGTC 300
Db 888 ATTGATTTTGCAGAGGATGCAACCTTATCTGATGCAAAACGGAAGTCTACTTTGTGAC 947
QY 301 ATTGATTTTGCAGAGGATGCAACCTTATCTGATGCAAAACGGAAGTCTACTTTGTGAC 360
Db 948 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 1007
QY 361 AAAGATGTTGGTAACTTGGCGAGCCCACTGTCCAGAAATTTTGTAAAGGACAT 420
Db 1008 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAACCAAGGAAATCTGTGATGTG 1067
QY 421 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAACCAAGGAAATCTGTGATGTG 480

```



```

QY 1068 GAATATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTGACACCAAGCTCAAA 1127
Db 481 GAAATATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTGACACCAAGCTCAAA 540
QY 1128 GATATCTAGCTATTTGAAAACATAGCTATCGGTGACACATCATCATCAAGATTTACTCA 1187
Db 541 GATATCTAGCTATTTGAAAACATAGCTATCGGTGACACATCATCATCAAGATTTACTCA 600
QY 1188 GCTCAGACCAAGAGATTTTAAACAAAACCAACCAAGGCTATACGATTTATGACGTGAC 1247
Db 601 GCTCAGACCAAGAGATTTTAAACAAAACCAACCAAGGCTATACGATTTATGACGTGAC 660
QY 1248 TCCCTCATCTGACATCATGACATGACATTTTCCGTGACATTTTACCAATGATCAAGAG 1307
Db 661 TCCCTCATCTGACATCATGACATGACATTTTCCGTGACATTTTACCAATGATCAAGAG 720
QY 1308 TTTACTTACCGGTGTTTAAATCCGGGACAGCTTATAGATCAATTAATTAATTAATCTGCTG 1367
Db 721 TTTACTTACCGGTGTTTAAATCCGGGACAGCTTATAGATCAATTAATTAATTAATCTGCTG 780
QY 1368 AATGAGAAATTAACAACTGACCTGATCTCTGAGAAATTAAGCTCTTAAAGAGAG 1427
Db 781 AATGAGAAATTAACAACTGACCTGATCTCTGAGAAATTAAGCTCTTAAAGAGAG 840
QY 1428 GAAAAGCCGTATGATCCCTTTGATGACGATCACTTGAACCTGTTCAACATCAATACGTT 1487
Db 841 GAAAAGCCGTATGATCCCTTTGATGACGATCACTTGAACCTGTTCAACATCAATACGTT 900
QY 1488 GATTCGATTCGAACGAAATTTGCTTAAAGTGAAGCTCTTAAAGCTTAAAGCTTAAAG 1547
Db 901 GATTCGATTCGAACGAAATTTGCTTAAAGTGAAGCTCTTAAAGCTTAAAGCTTAAAG 960
QY 1548 TTAGACTTCAGAGATTTTATAGATCTCTGATGATGATGATGATGATGATGATGATGATG 1607
Db 961 TTAGACTTCAGAGATTTTATAGATCTCTGATGATGATGATGATGATGATGATGATGATG 1020
QY 1608 GATGCTTTGGTATTTATGACCTATACCTTAACTGGAAGAAATGAGATATACAGATGAC 1667
Db 1021 GATGCTTTGGTATTTATGACCTATACCTTAACTGGAAGAAATGAGATATACAGATGAC 1080
QY 1668 ACCAAGCGTATCATACCGTTTATATGAGGCAAGCGACCCGAGAGAGAGAAATGCTATAC 1727
Db 1081 ACCAAGCGTATCATACCGTTTATATGAGGCAAGCGACCCGAGAGAGAGAAATGCTATAC 1140
QY 1728 CATTTAGCTGTGCT 1742
Db 1141 CATTTAGCTGTATGAT 1155

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
;
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-488-940-19

```

Query Match 54.0%; Score 1132.4; DB 2; Length 2566;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1148; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

```

QY 587 TAATGCTGACCTGATGCTGCTGATGACCTGCTCATCTGTCACAAACAGCCAAATGGTGT 646
Db 896 TAATGCTGACCTGATGCTGCTGCTGATGACCTGCTCATCTGTCACAAACAGCCAAATGGTGT 955
QY 647 TAGCGTGTGGGATCTGTTGAGGGGAGATCAAGATTAAGCTTAATTTTGAAT 706
Db 956 TAGCGTGTGGGATCTGTTGAGGGGAGATCAAGATTAAGCTTAATTTTGAAT 1015
QY 707 CGATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGAGCTTAAGTCCAAATC 766
Db 1016 CGATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGAGCTTAAGTCCAAATC 1073
QY 767 AAAACATTTGCTACTGATGATGCGCGGATGTCATTAACCTTGAGAAAGCTGACTTACT 826
Db 1074 AAAACATTTGCTACTGATGATGCGCGGATGTCATTAACCTTGAGAAAGCTGACTTACT 1133
QY 827 AAAGGCTATTCAGAAACAAATGATGCTTAACGTCACATTAACGACACTTCTGAGGT 886
Db 1134 AAAGGCTATTCAGAAACAAATGATGCTTAACGTCACATTAACGACACTTCTGAGGT 1193
QY 887 CATTAATTTGCAAGCGATGCAACATTAATGATGCAAAAGGCAAGGTCTACTTTGCTGA 946
Db 1194 CATTAATTTGCAAGCGATGCAACATTAATGATGCAAAAGGCAAGGTCTACTTTGCTGA 1253
QY 947 CAAAGATGTTGCTGATACCTTGCCGACCCGACCTGTCAGAAATTTTTCAGGCGACA 1006
Db 1254 CAAAGATGTTGCTGATACCTTGCCGACCCGACCTGTCAGAAATTTTTCAGGCGACA 1313
QY 1007 TGTGGCGCTTACACATTAAGAAACCAATATACAAACCAAGGCAATCTGTTATGCT 1066
Db 1314 TGTGGCGCTTACACATTAAGAAACCAATATACAAACCAAGGCAATCTGTTATGCT 1373
QY 1067 GGAATATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTGACAGCGGTCAA 1126
Db 1374 GGAATATATCTGACAGTTTACTCCCTTAACCTGATGACGATTTGACAGCGGTCAA 1433
QY 1127 AGATATCTAGCTATTTGAAAACATAGCTATCGGTGACACATCATCATCAAGATTTACT 1186
Db 1434 AGATATCTAGCTATTTGAAAACATAGCTATCGGTGACACATCATCATCAAGATTTACT 1493
QY 1187 AGCTCAAGACCAAGAGATTTTAAACAAAACCAACCAAGGCTATACGATTTATGAACTGA 1246
Db 1494 AGCTCAAGACCAAGAGATTTTAAACAAAACCAACCAAGGCTATACGATTTATGAACTGA 1553
QY 1247 CTCCTCAATCGCTCATGATGACATGACATTTTCCGTGACATTTTACCAATGATCAAGAG 1306
Db 1554 CTCCTCAATCGCTCATGATGACATGACATTTTCCGTGACATTTTACCAATGATCAAGAG 1613
QY 1307 GTTTACTTACCGTGTAAATAATCGGGAACAACTTATAGATCAATTAATAATCTGCTGT 1366
Db 1614 GTTTACTTACCGTGTAAATAATCGGGAACAACTTATAGATCAATTAATAATCTGCTGT 1673

```

```

OY 1367 GAATGAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1426
    |||||
DB 1674 GAATGAGAAATAAACAACACGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1733
    |||||
OY 1427 GAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTACCATCAAAATACGT 1486
    |||||
DB 1734 GAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTACCATCAAAATACGT 1793
    |||||
OY 1487 TGATGTCATACCAACGAATTTGCTAAAAAGTAGCAGCTCTTAAACAGTAGCAAGCTAA 1546
    |||||
DB 1794 TGATGTCATACCAACGAATTTGCTAAAAAGTAGCAGCTCTTAAACAGTAGCAAGCTAA 1853
    |||||
OY 1547 CTTAGACTTCAAGATTATATACGATCTCTGATAGGCTTAACTACTCTTCAACAACATCT 1606
    |||||
DB 1854 CTTAGACTTCAAGATTATATACGATCTCTGATAGGCTTAACTACTCTTCAACAACATCT 1913
    |||||
OY 1607 CGATGCTTTGTTATTTATGACCTATACCTTAACTGAAAAAGTAGAGATAATCAAGATGA 1666
    |||||
DB 1914 CGATGCTTTGTTATTTATGACCTATACCTTAACTGAAAAAGTAGAGATAATCAAGATGA 1973
    |||||
OY 1667 CACCAACCGTATATACCCGTTTATATGGGCAAGCCGCAAGAGAGAAATGCTTACGTA 1726
    |||||
DB 1974 CACCAACCGTATATACCCGTTTATATGGGCAAGCCGCAAGAGAGAAATGCTTACGTA 2033
    |||||
OY 1727 CCATTTAGCTGTGTGT 1742
    |||||
DB 2034 TCATTTAGCCTATGAT 2049
    |||||

```

RESULT 6

```

US-07-854-596B-42
; Sequence 42, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1458 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1458

```

```

; OTHER INFORMATION: /note= "Hirudin-streptokinase
; OTHER INFORMATION: fusion linked by Factor Xa cleavable IGR"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1449
; FEATURE:
; NAME/KEY: mac_peptide
; LOCATION: 1..1449
; US-07-854-596B-42

```

```

Query Match 53.7%; Score 1125.4; DB 2; Length 1458;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1141; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

```

OY 576 ACCGATGTTGCTATTGCTGGAACCTGAGTGGCGCTAGAACCCGTCATCTGTCAACAACGCC 635
    |||||
DB 136 ATCGAAGGTAAATATGCTGACCTTAGTGGCTGTAGACCGTCACTGTCAACAACGCC 255
    |||||
OY 636 CAATTGCTTTAGCGTTGCTGTGTATCTTTGAGGGAAGCAATCAAGACATTAGCTTTAA 695
    |||||
DB 256 CAATTAGTTGTTAGCGTTGCTGTGTATCTTTGAGGGAAGCAATCAAGACATTAGCTTTAA 315
    |||||
OY 696 TTTTGTGAAATCGATCTTAAATCAATCAAGCTGTCTCATGAGAGAAAGACAGAGAGCTTTA 755
    |||||
DB 316 TTTTGTGAAATCGATCTTAAATCAATCAAGCTGTCTCATGAGAGAAAGACAGAGAGCTTTA 375
    |||||
OY 756 AGTCGAAATGAAACCATTTGCTATCTATAGTGGCGGATGTCATTTAACTTATAGAA 815
    |||||
DB 376 AGTCGAAATGAAACCATTTGCTATCTATAGTGGCGGATGTCATTTAACTTATAGAA 435
    |||||
OY 816 GCTGACTTACTTAAAGGCTATTCAAGAACATTTGATGCTTAACTGCTCAAGTAAAGACGAC 875
    |||||
DB 436 GCTGACTTACTTAAAGGCTATTCAAGAACATTTGATGCTTAACTGCTCAAGTAAAGACGAC 495
    |||||
OY 876 TACTTTGAGTCAATGATTTTTCGAAAGCATGCAACATTACTGATCGAAACGGCAAGTTC 935
    |||||
DB 496 TACTTTGAGTCAATGATTTTTCGAAAGCATGCAACATTACTGATCGAAACGGCAAGTTC 555
    |||||
OY 936 TACTTTGCTGACAAAGATGTTGCTGTAACCTTGGCCGACCACTGTCCAAATTTTGG 995
    |||||
DB 556 TACTTTGCTGACAAAGATGTTGCTGTAACCTTGGCCGACCACTGTCCAAATTTTGG 615
    |||||
OY 996 CTAAGCGGACATGTCGGGCTTATGACCATATTAAGAAAAACCAATCAAAACCAAGGAA 1055
    |||||
DB 616 CTAAGCGGACATGTCGGGCTTATGACCATATTAAGAAAAACCAATCAAAACCAAGGAA 675
    |||||
OY 1056 TCTGTTGATGTAATATATCTGACATTTACTCCCTTAAACCTGATGACGATTTTCA 1115
    |||||
DB 676 TCTGTTGATGTAATATATCTGACATTTACTCCCTTAAACCTGATGACGATTTTCA 735
    |||||
OY 1116 CCAGGCTCAAAAGATCTAGCTATTTGAAAAACATGACTATGCTGACCATTCATCTT 1175
    |||||
DB 736 CCAGGCTCAAAAGATCTAGCTATTTGAAAAACATGACTATGCTGACCATTCATCTT 795
    |||||
OY 1176 CAAGAATTACTAGCTCAAGCACAAGCACTTTTAAACAAACCAACCCGAGCTATGAT 1235
    |||||
DB 796 CAAGAATTACTAGCTCAAGCACAAGCACTTTTAAACAAACCAACCCGAGCTATGAT 855
    |||||
OY 1236 TATGAACGTGACTCCTCATATGCTCACTCATGACATTTTCCGTACGATTTTACCA 1295
    |||||
DB 856 TATGAACGTGACTCCTCATATGCTCACTCATGACATTTTCCGTACGATTTTACCA 915
    |||||
OY 1296 ATGATCAAGAGTTTACTTACCTGTTAAAAATCGGAAACAAGCTTATAGATCAATATA 1355
    |||||
DB 916 ATGATCAAGAGTTTACTTACCTGTTAAAAATCGGAAACAAGCTTATAGATCAATATA 975
    |||||
OY 1356 AAATCTGCTGATGAGAAATAAACAACACTGACCTGATCTTTCGAGAAATATTACGTC 1415
    |||||
DB 976 AAATCTGCTGATGAGAAATAAACAACACTGACCTGATCTTTCGAGAAATATTACGTC 1035
    |||||
OY 1416 CTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTACCC 1475
    |||||
DB 1036 CTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAAACTGTTACCC 1095
    |||||

```

QY 1476 ATCAATAGTGTGATGTCGATACCAACGAATGCTTAAAGAGAGAGCTTTAAACAGCT 1535
Db 1096 ATCAATAGTGTGATGTCGATACCAACGAATGCTTAAAGAGAGAGCTTTAAACAGCT 1155
QY 1536 AGCGAAGCTTACCTTGAAGCTTCAAGATTTATATACGATCTCTGATTAAGGCTTAACTATCTC 1595
Db 1156 AGCGAAGCTTACCTTGAAGCTTCAAGATTTATATACGATCTCTGATTAAGGCTTAACTATCTC 1215
QY 1596 TACAACATCTCGAGCTTTGGTATATGAGCTATACCTTAAGTGAAGAGAT 1655
Db 1216 TACAACATCTCGAGCTTTGGTATATGAGCTATACCTTAAGTGAAGAGAT 1275
QY 1656 AATCAGATGACCAACCGTATCATTAACGTTTATATGAGGCAAGGAGCCGGAAGAGAG 1715
Db 1276 AATCAGATGACCAACCGTATCATTAACGTTTATATGAGGCAAGGAGCCGGAAGAGAG 1335
QY 1716 AATGCTAGCTACCATTTAGCTGATGCT 1742
Db 1336 AATGCTAGCTACCATTTAGCTGATGCT 1362

RESULT 7
US-07-703-778D-1
Sequence 1, Application US/07703778D
Patent No. 5296366
GENERAL INFORMATION:
APPLICANT: Garcia, M.P.E. et al
TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
TITLE OF INVENTION: OF A GENE WHICH CODES FOR STREPTOKINASE, NUCLEOTIDE SEQUENCE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stenger, Michaelson, Spivak and Wallace, Esq.
STREET: Parkway 109 Office Center, 328 Newman Springs Road,
CITY: Red Bank
STATE: New Jersey
COUNTRY: USA
ZIP: 07701

COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2" 1.44mb IBM compatible diskette
COMPUTER: IBM PS/2 Model 80
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Microsoft Word for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07703,778D
FILING DATE: 19910522
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 30090
REFERENCE/DOCKET NUMBER: Centro-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 530-6671
FAX: (908) 530-6584
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1245 base pairs
TYPE: NUCLEOTIDE WITH CORRESPONDING PROTEIN
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Streptococcus equisimilis from group C of Lanfield
ORGANISM: definition
IMMEDIATE SOURCE: ATCC-9542 strain
FEATURE: from 1 to 1245 bp mature peptide
OTHER INFORMATION:
OTHER INFORMATION: Properties: Streptokinase gene
OTHER INFORMATION: The gene product is an activator of human plasminogen
OTHER INFORMATION: The gene product is an activator of human plasminogen
US-07-703-778D-1

Query Match 53.7%; Score 1124.6; DB 2; Length 1245;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 588 ATTGCTGAGACCTGAGTGGCTGCTAGACGTCATCTGTCAACCAAGCCAAATGGTGTG 647
Db 1 ATTGCTGAGACCTGAGTGGCTGCTAGACGTCATCTGTCAACCAAGCCAAATGGTGTG 60
QY 648 AGCGTCTGCTGATCTGTGAGGGGAGCAATCAAGCATTAAGCTTTAAATTTTGAATTC 707
Db 61 AGCGTCTGCTGATCTGTGAGGGGAGCAATCAAGCATTAAGCTTTAAATTTTGAATTC 120
QY 708 GATCTAACATCAGACCGTGCATGAGGAAGCAGAGCAAGGCTTAAGTCAAAAATCA 767
Db 121 GATCTAACATCAGACCGTGCATGAGGAAGCAGAGCAAGGCTTAAGTCAAAAATCA 180
QY 768 AAACCATTTGCTACTGATAGTGGCGCATGTCACATTAACCTTGAGAAAGCTGACTTACTA 827
Db 181 AAACCATTTGCTACTGATAGTGGCGCATGTCACATTAACCTTGAGAAAGCTGACTTACTA 240
QY 828 AAGGCTATTCAGAAACATTTGATGCTTAACGTCACAGTAACGACGACTTATGAGGTC 887
Db 241 AAGGCTATTCAGAAACATTTGATGCTTAACGTCACAGTAACGACGACTTATGAGGTC 300
QY 888 ATTGATTTTCCAGGCGATGCAACCATTAAGTGAAGAAAGGCAAGGCTTATGCTGAC 947
Db 301 ATTGATTTTCCAGGCGATGCAACCATTAAGTGAAGAAAGGCAAGGCTTATGCTGAC 360
QY 948 AAAGATGTTTGGTAACTTGGCGGACCACTGTCGAAGAAATTTTGGTAAAGGAGCAT 1007
Db 361 AAAGATGTTTGGTAACTTGGCGGACCACTGTCGAAGAAATTTTGGTAAAGGAGCAT 420
QY 1008 GTGGCGCTTGAACCATTAAGAAAGAAACCAATACAAACCAAGGAAATCTGTGATGTG 1067
Db 421 GTGGCGCTTGAACCATTAAGAAAGAAACCAATACAAACCAAGGAAATCTGTGATGTG 480
QY 1068 GAATATATCTGATGATTTACTCTCTTAAACCTGATGACATTTTCAAGCCAGGCTCAMA 1127
Db 481 GAATATATCTGATGATTTACTCTCTTAAACCTGATGACATTTTCAAGCCAGGCTCAMA 540
QY 1128 GATACCTAAGCTATTTGAACCACTAGCTATGAGTGAACCAATCACTGCAAGAAATTTCTA 1187
Db 541 GATACCTAAGCTATTTGAACCACTAGCTATGAGTGAACCAATCACTGCAAGAAATTTCTA 600
QY 1188 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTGAAGCTGAC 1247
Db 601 GCTCAAGCACAAGCATTTTAAACAAACCAACCCAGGCTATACGATTTTGAAGCTGAC 660
QY 1248 TCCTCAATGCTCACTGACATGACATGACATTTTCCGTAACGATTTTCAATGATCAAGAG 1307
Db 661 TCCTCAATGCTCACTGACATGACATGACATTTTCCGTAACGATTTTCAATGATCAAGAG 720
QY 1308 TTACTTACCGGTGTTAAATTCGGGAAACAGCTTATAGATCAATTAATAATCGGTCTG 1367
Db 721 TTACTTACCGGTGTTAAATTCGGGAAACAGCTTATAGATCAATTAATAATCGGTCTG 780
QY 1368 AATGAAGAATAAACAACCTGACCTGATCTGAGAAATATTTCTGTTAAAGAGG 1427
Db 781 AATGAAGAATAAACAACCTGACCTGATCTGAGAAATATTTCTGTTAAAGAGG 840
QY 1428 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAACGTGTTCAACATCAATACGTT 1487
Db 841 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAACGTGTTCAACATCAATACGTT 900
QY 1488 GATGTGATTCAGACGAATTTGCTTAAAGAGGAGAGCTTTTACAGCTAGGAGGATAC 1547
Db 901 GATGTGATTCAGACGAATTTGCTTAAAGAGGAGAGCTTTTACAGCTAGGAGGATAC 960
QY 1548 TTAGCTTGAAGATTTTATGATCCGATTAAGGCTTAACTACTCTCAACATCTC 1607
Db 961 TTAGCTTGAAGATTTTATGATCTCTGATTAAGGCTTAACTACTCTCAACATCTC 1020

APPLICANT: Czaplowski, Lloyd G
 TITLE OF INVENTION: Proteins and nucleic acids
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dr. John J. McDonnell
 STREET: Ten South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B
 FILING DATE: 03-JUN-1992
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949
 REFERENCE/DOCKET NUMBER: 92,337
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 910-221-5317
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1257 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..1257
 OTHER INFORMATION: /note="Methionyl-streptokinase"
 OTHER INFORMATION: fusion protein"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4..1248
 NAME/KEY: mat peptide
 LOCATION: 4..1248
 US-07-854-596B-25

Query Match 53.5%; Score 1121.4; DB 2; Length 1257;
 Best Local Similarity 98.2%; Pred. No. 0;
 Matches 1134; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 588 ATTGCTGACCTGAGTGGCTGCTGACCGTTCATCTGTCAACACAGCCAAATGTTGTT 647
 DB 7 ATTGCTGACCTGAGTGGCTGCTGACCGTTCATCTGTCAACACAGCCAAATGTTGTT 66
 QY 648 AGCGTTGCTGTACTGTGAGGGGACGAATCAAGACATTGATTTTGAATTC 707
 DB 67 AGCGTTGCTGTACTGTGAGGGGACGAATCAAGACATTGATTTTGAATTC 126
 QY 708 GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAGTCAAAATCA 767
 DB 127 GACCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAGTCAAAATCA 186
 QY 768 AAACCATTTGCTATGATAGTGGCGCGATGTCATTAATCTTGAGAAAGCTGACTTACTA 827
 DB 187 AAACCATTTGCTATGATAGTGGCGCGATGTCATTAATCTTGAGAAAGCTGACTTACTA 246
 QY 828 AAGGCTATTCAAGAAATGATGCTGTAAGTCAAGTCAAGCACTTCTTTAGAGTC 887
 DB 247 AAGGCTATTCAAGAAATGATGCTGTAAGTCAAGTCAAGCACTTCTTTAGAGTC 306
 QY 888 ATTGATTTTGAAGGATGCAACATTAATGATGAAAGGCAAGGCTACTTTGCTGAC 947

DB 307 ATTGATTTTGAAGGATGCAACATTAATGATGAAAGGCAAGGCTACTTTGCTGAC 366
 QY 948 AAAGATGTTGGTAACTTGGCCGACCCAACTGTCAGAAATTTTGTAGCCGACAT 1007
 DB 367 AAAGATGTTGGTAACTTGGCCGACCCAACTGTCAGAAATTTTGTAGCCGACAT 426
 QY 1008 GTGCGCTTGAACCATATTAAGAAACCAATACAAACCAAGCAATCTGTTAGTGT 1067
 DB 427 GTGCGCTTGAACCATATTAAGAAACCAATACAAACCAAGCAATCTGTTAGTGT 486
 QY 1068 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTGAACAGTCTCAA 1127
 DB 487 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTGAACAGTCTCAA 546
 QY 1128 GATATAGCTATTGAAACCACTACTCTGCTGACACATCACTTCAAGAAATTTACTA 1187
 DB 547 GATATAGCTATTGAAACCACTACTCTGCTGACACATCACTTCAAGAAATTTACTA 606
 QY 1188 GCTCAAGCACAAGCAATTTTAAACAAACCAAGGCTATACGATTTTGAACAGTGCAC 1247
 DB 607 GCTCAAGCACAAGCAATTTTAAACAAACCAAGGCTATACGATTTTGAACAGTGCAC 666
 QY 1248 TCTCAATGCTCACTCATGACATGACATTTTCCGTAGCAATTTTCAATGATCAAGAG 1307
 DB 667 TCTCAATGCTCACTCATGACATGACATTTTCCGTAGCAATTTTCAATGATCAAGAG 726
 QY 1308 TTTACTTACCGTGTAAATTCGGGACCAAGCTTTAGATTAATTAATCTGCTGTG 1367
 DB 727 TTTACTTACCGTGTAAATTCGGGACCAAGCTTTAGATTAATTAATCTGCTGTG 786
 QY 1368 AATGAAGAAATTAACCAACTGACCTGATCTGAGAAATTTAGTCTTTAAAGAGG 1427
 DB 787 AATGAAGAAATTAACCAACTGACCTGATCTGAGAAATTTAGTCTTTAAAGAGG 846
 QY 1428 GAAAAACCGTATGATCCCTTGAATGCACTGACCTTGAATCTGTTCAACATCAATACGTT 1487
 DB 847 GAAAAACCGTATGATCCCTTGAATGCACTGACCTTGAATCTGTTCAACATCAATACGTT 906
 QY 1488 GATGTCATACCAAGAAATGCTTAAGAAAGTGAAGCTCTTACAGTACGAAAGTAC 1547
 DB 907 GATGTCATACCAAGAAATGCTTAAGAAAGTGAAGCTCTTACAGTACGAAAGTAC 966
 QY 1548 TTAGACTTCAAGAAATTAATGATGCTGCTGATTAAGCTTAAGCTTCAACAAATCTC 1607
 DB 967 TTAGACTTCAAGAAATTAATGATGCTGCTGATTAAGCTTAAGCTTCAACAAATCTC 1026
 QY 1608 GATGCTTTGCTATTAAGAACTATACCTTAAGTGAAGAAAGTGAAGATTAACAGATAC 1667
 DB 1027 GATGCTTTGCTATTAAGAACTATACCTTAAGTGAAGAAAGTGAAGATTAACAGATAC 1086
 QY 1668 ACCAACCGTATCATACCGTTTATATGAGGCAAGGACCCGAAGAGAGAAATGCTTAC 1727
 DB 1087 ACCAACCGTATCATACCGTTTATATGAGGCAAGGACCCGAAGAGAGAAATGCTTAC 1146
 QY 1728 CATTAGCTGTGCT 1742
 DB 1147 CATTAGCTGTGAT 1161

RESULT 12
 US-07-854-596B-18
 ; Sequence 18, Application US/07854596B
 ; Patent No. 5434073
 ; GENERAL INFORMATION:
 ; APPLICANT: Dawson, Keith M
 ; APPLICANT: Hunter, Michael G
 ; APPLICANT: Czaplowski, Lloyd G
 ; TITLE OF INVENTION: Proteins and nucleic acids
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dr. John J. McDonnell
 ; STREET: Ten South Wacker Drive, Suite 3000
 ; CITY: Chicago

STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDowell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1317 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1317
OTHER INFORMATION: /note="OmpA fused to mature
FBATURE: streptokinase gene"
NAME/KEY: CDS
LOCATION: 4..1308
FBATURE:
NAME/KEY: mat peptide
LOCATION: 4..1308
US-07-854-596B-18

Query Match 53.5%; Score 1121.4; DB 2; Length 1317;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 588 ATTGCTGACCTGAGTGGTCTGCTAGACCGTCCATCTGTCAACAACGCCAATTGGTTGTT 647
DB 67 ATTGCTGACCTGAGTGGTCTGCTAGACCGTCCATCTGTCAACAACGCCAATTGGTTGTT 126
QY 648 AGCGTTGCTGACTGTTGAGGGGAGCAATCAAGATCAATTAGTCTTAATTTTGAATC 707
DB 127 AGCGTTGCTGACTGTTGAGGGGAGCAATCAAGATCAATTAGTCTTAATTTTGAATC 186
QY 708 GATCTACATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 767
DB 187 GACCTAACATCAACGACCTGCTCATGAGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 246
QY 768 AACCACTTGTCTACTGATAGTGGCGCGATGTCACATPAACTTGAGAAAGTCACTTACTA 827
DB 247 AACCACTTGTCTACTGATAGTGGCGCGATGTCACATPAACTTGAGAAAGTCACTTACTA 306
QY 828 AAGCTATTCAAGAACATGATTCGCTAAGTCCAGATGACAGACTACTTTGAGGTC 887
DB 307 AAGCTATTCAAGAACATGATTCGCTAAGTCCAGATGACAGACTACTTTGAGGTC 366
QY 888 ATTGATTTTGGCAAGGATGCAACATTAATGATCGAAGCGCAAGGCTTACTTGGTGAC 947
DB 367 ATTGATTTTGGCAAGGATGCAACATTAATGATCGAAGCGCAAGGCTTACTTGGTGAC 426
QY 948 AAAAGATGTTGGTAACCTTGGCGACCAACTGCTGCAAGAAATTTTGGTAAAGCGACAT 1007
DB 427 AAAAGATGTTGGTAACCTTGGCGACCAACTGCTGCAAGAAATTTTGGTAAAGCGACAT 486
QY 1008 GTGGCGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTC 1067

|||||
DB 487 GTGGCGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTC 546
QY 1068 GAATATCTGTACAGTTTACTCCCTTAAACCGTAGACATTTTCAACCGGCTCAAA 1127
DB 547 GAATATCTGTACAGTTTACTCCCTTAAACCGTAGACATTTTCAACCGGCTCAAA 606
QY 1128 GATPCTAAGCTATTGAAAACACTAGTATGAGTGAACCAATCATCATCTCAAGATTACTA 1187
DB 607 GATPCTAAGCTATTGAAAACACTAGTATGAGTGAACCAATCATCATCTCAAGATTACTA 666
QY 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGAGCTTAAGATTATGAACGTGAC 1247
DB 667 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGAGCTTAAGATTATGAACGTGAC 726
QY 1248 TCCCTCAATCGTCACTGACATGACATGATTTTCCGTACGATTTTCAACATGATCAAG 1307
DB 727 TCCCTCAATCGTCACTGACATGACATGATTTTCCGTACGATTTTCAACATGATCAAG 786
QY 1308 TTTACTTACCGTGTAAATAATCGGGAACAAAGCTTATGATCAATTAATAATCTGCTG 1367
DB 787 TTTACTTACCGTGTAAATAATCGGGAACAAAGCTTATGATCAATTAATAATCTGCTG 846
QY 1368 AATGAAGAAATTAACAACAACCTGACCTGATCTGAGAAATATTAAGTCTTAAAAAGG 1427
DB 847 AATGAAGAAATTAACAACAACCTGACCTGATCTGAGAAATATTAAGTCTTAAAAAGG 906
QY 1428 GAAAAGCGTATGATCCCTTTGATCGCATCTTGAACCTGTTCAACATCAATTAAGTT 1487
DB 907 GAAAAGCGTATGATCCCTTTGATCGCATCTTGAACCTGTTCAACATCAATTAAGTT 966
QY 1488 GATGTGATPCCAAGCAATGCTTAABAAGAGAGAGCTTAAAGCTAGGGAACGTAC 1547
DB 967 GATGTCAACCAAGCAATGCTTAABAAGAGAGAGCTTAAAGCTAGGGAACGTAC 1026
QY 1548 TTAGACTTCAGAGATTATACATCTCGATPAGGCTTAACTCTCAACATCTC 1607
DB 1027 TTAGACTTCAGAGATTATACATCTCGATPAGGCTTAACTCTCAACATCTC 1086
QY 1608 GATGCTTTGATATTAAGACTATACCTTAACCTGAAAAGTAGAGATTAACAGATGAC 1667
DB 1087 GATGCTTTGATATTAAGACTATACCTTAACCTGAAAAGTAGAGATTAACAGATGAC 1146
QY 1668 ACCAAGCTATCAATACCGTTTATATGAGCAAGCAACCGAAGAGAAATGCTAGTAC 1727
DB 1147 ACCAAGCTATCAATACCGTTTATATGAGCAAGCAACCGAAGAGAAATGCTAGTAT 1206
QY 1728 CATTAGCTGTGAT 1742
DB 1207 CATTAGCTGTGAT 1221
RESULT 13
US-07-854-596B-46
Sequence 46, Application US/07854596B
Patent No. 5434073
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Czaplowski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSER: Dr. John J. McDowell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1467 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1467
OTHER INFORMATION: /note= "Streptokinase-hirudin
OTHER INFORMATION: fusion linked by Factor Xa-cleavable IEGF"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1449
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..1449
US-07-854-596B-46

Query Match 53.5%; Score 1121.4; DB 2; Length 1467;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 588 ATTGCTGACCTGAGTGGCTGCTGACCGCTCATCTGTCAACAACGCCAATTGGTTGT 647
DB 1 ATTGCTGACCTGAGTGGCTGCTGACCGCTCATCTGTCAACAACGCCAATTGGTTGT 60
QY 648 AGCGTTGCTGTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
DB 61 AGCGTTGCTGTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 120
QY 708 GATTTAACATCAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
DB 121 GACCTTAACATCAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 180
QY 768 AAACGATTGCTAGTGTAGTGGCGGAGTGTCAATAAATTGAGAAAGCTGACTTA 827
DB 181 AAACGATTGCTAGTGTAGTGGCGGAGTGTCAATAAATTGAGAAAGCTGACTTA 240
QY 828 AAGGCTATTCAGAAACAATTGATCGTAAAGTCCAGACGTAAGCACTACTTTGAGGTC 887
DB 241 AAGGCTATTCAGAAACAATTGATCGTAAAGTCCAGACGTAAGCACTACTTTGAGGTC 300
QY 888 ATTGATTGTCAGAGCGATGCAACATTACTGATGAAACGGCAAGGCTTACTTTGCTGAC 947
DB 301 ATTGATTGTCAGAGCGATGCAACATTACTGATGAAACGGCAAGGCTTACTTTGCTGAC 360
QY 948 AAAGATGTTGGGTAACTTTGCCGACCACTGCTCCAGAAATTTTGGTAAAGCGACAT 1007
DB 361 AAAGATGTTGGGTAACTTTGCCGACCACTGCTCCAGAAATTTTGGTAAAGCGACAT 420
QY 1008 GTGCGGTGAGACCATATAAGAAAAACAATAAACAACGAAATGTTGATGAG 1067
DB 421 GTGCGGTGAGACCATATAAGAAAAACAATAAACAACGAAATGTTGATGAG 480
QY 1068 GAATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTGAGACCAAGGCTCAAA 1127
DB 481 GAATATCTGTAAGTTTACTCCCTTAAACCTGATGACGATTTGAGACCAAGGCTCAAA 540

QY 1128 GATATCAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1187
DB 541 GATATCAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600
QY 1188 GCTAAGACCAAGCAATTTTAAACAAAACCAACCGGCTATAGCATTTATGAACGTGAC 1247
DB 601 GCTAAGACCAAGCAATTTTAAACAAAACCAATCAGGCTATAGCATTTATGAACGTGAC 660
QY 1248 TCCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTAAACCAATGATCAAG 1307
DB 661 TCCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTAAACCAATGATCAAG 720
QY 1308 TTACTTACCGTGTAAAAAATCGGAAACAAGCTTAAGATCAATTAATAAATCGTCTG 1367
DB 721 TTACTTACCGTGTAAAAAATCGGAAACAAGCTTAAGATCAATTAATAAATCGTCTG 780
QY 1368 AATGAGAAATTAACAAACACTGACCTGATCTGTGAGAAATTAAGTCTTTAAAAAGG 1427
DB 781 AATGAGAAATTAACAAACACTGACCTGATCTGTGAGAAATTAAGTCTTTAAAAAGG 840
QY 1428 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAACCTTCAACATATACGT 1487
DB 841 GAAAAGCGTATGATCCCTTTGATCGAGTCACTTGAACCTTCAACATATACGT 900
QY 1488 GATGCTGATCAACGAATGCTTAAAAAGTAGAGAGCTTTAAAGAGTGAAGCGTAAC 1547
DB 901 GATGCTGATCAACGAATGCTTAAAAAGTAGAGAGCTTTAAAGAGTGAAGCGTAAC 960
QY 1548 TTAGACTTCAGAGATTTATAGATCTGTGATTAAGGCTTAACCTTCAACATCTC 1607
DB 961 TTAGACTTCAGAGATTTATAGATCTGTGATTAAGGCTTAACCTTCAACATCTC 1020
QY 1608 GATGCTTTGGTATTTATAGACTATACCTTAACTGAAAAAGTAGAGATTAACAGATGAC 1667
DB 1021 GATGCTTTGGTATTTATAGACTATACCTTAACTGAAAAAGTAGAGATTAACAGATGAC 1080
QY 1668 ACCAAGCGTATCAATAACGTTTATATAGGAGGAGACCCGAAAGAGAAATGCTAGTAC 1727
DB 1081 ACCAAGCGTATCAATAACGTTTATATAGGAGGAGACCCGAAAGAGAAATGCTAGTAC 1140
QY 1728 CATTAGCTGTGCT 1742
DB 1141 CATTAGCTGTGCT 1155

RESULT 14
US-09-374-038-13
Sequence 13, Application US/09374038
Patent No. 6309873
GENERAL INFORMATION:
APPLICANT: Madrazo, Isis Del Carmen Torrens
APPLICANT: Garcia, Jose De Jesus De la Fuente
APPLICANT: Ojalvo, Ariana Garcia
APPLICANT: Menendez, Alina Seralena
APPLICANT: Escalona, Elder Pupo
APPLICANT: Masso, Julio Raul Fernandez
APPLICANT: Griego, Martha De Jesus Gonzalez
TITLE OF INVENTION: STREPTOKINASE MUTANTS
FILE REFERENCE: Sequence listings 1-14 re: 976-5
Patent No. 6309873
CURRENT APPLICATION NUMBER: US/09/374,038
CURRENT FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 1122
TYPE: DNA
ORGANISM: Streptococcus equisimilis
US-09-374-038-13

Query Match 52.2%; Score 1093.6; DB 3; Length 1122;
Best Local Similarity 98.7%; Pred. No. 3,66-312;
Matches 1102; Conservative 0; Mismatches 14; Indels 0; Gaps 0;


```
QY 1188 GCTCAAGCACAAAGCATTTTAAACAAAACCAAGGCTATACGATTATGAAAGTGAC 1247
Db 604 GCTCAAGCACAAAGCATTTTAAACAAAACCAAGGCTATACGATTATGAAAGTGAC 663
QY 1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTACCAATGATCAAGAG 1307
Db 664 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTACCAATGATCAAGAG 723
QY 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATPAAAAAATCTGGTCTG 1367
Db 724 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATPAAAAAATCTGGTCTG 783
QY 1368 AATGAAGAAAATAAACCACTGACCTGATCTGTGAAATATTAAGTCTTAAAAAAGG 1427
Db 784 AATGAAGAAAATAAACCACTGACCTGATCTGTGAAATATTAAGTCTTAAAAAAGG 843
QY 1428 GAAAAGCCGTATGATCCCTTTGATGCGATGCACTTGAAACTGTTCAACCATCAAAATACGTT 1487
Db 844 GAAAAGCCGTATGATCCCTTTGATGCGATGCACTTGAAACTGTTCAACCATCAAAATACGTT 903
QY 1488 GATGTGATACCAAGAAATGCTAAAGAGAGAGAGCTCTTAACAGCTAAGGAAGCTAAC 1547
Db 904 GATGTGATACCAAGAAATGCTAAAGAGAGAGAGCTCTTAACAGCTAAGGAAGCTAAC 963
QY 1548 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACACACATCTC 1607
Db 964 TTAGACTTCAGAGATTTATACGATCTCTGATGATAAGGCTAAACTACTCTACACACATCTC 1023
QY 1608 GATGCTTTTGGTATATGACATACTTAACCTGAAAGTAGAGATTAATCAGATGAC 1667
Db 1024 GATGCTTTTGGTATATGACATACTTAACCTGAAAGTAGAGATTAATCAGATGAC 1083
QY 1668 ACCAACCCTATCAACCGTTTATATGAGGCAAGCGA 1703
Db 1084 ACCAACCCTATCAACCGTTTATATGAGGCAAGCGA 1119
```

Search completed: February 1, 2006, 12:43:44
Job time : 371.056 secs


```

Db      121  GGGTCCCTCAACGACAGAGACGATCATGCGCACCCGTGCGCAGAACCCGATGCGCG 180
Qy      181  AGATCTCGATCCCGCGCAATTAATACGACTCATATGGGAGACCAACGGTTCCCTC 240
Db      181  |||||
Db      181  AAGTCTCGATCCCGCAAAATTAATACGACTCATATGGGAGACCAACGGTTCCCTC 240
Qy      241  TGAATAATATTTTGTAACTTTAAGAGAGATATACATGCTGTCAGGACCAACAGATT 300
Db      241  TGAATAATATTTTGTAACTTTAAGAGAGATATACATGCTGTCAGGACCAACAGATT 300
Qy      301  GTAACCATAGCTGAGAGAGTGTGTTGATCATGCTGCTGGGACTTCTATGCTGAGAA 360
Db      301  GTAACCATAGCTGAGAGAGTGTGTTGATCATGCTGCTGGGACTTCTATGCTGAGAA 360
Qy      361  AGGTGGAGAGAGGACCGGACCGCATCATCTTGCACTTTAGAAATGATGCAACGATCAGG 420
Db      361  AGGTGGAGAGAGGACCGGACCGCATCATCTTGCACTTTAGAAATGATGCAACGATCAGG 420
Qy      421  ACACAGAGCATCTCTATAGAAATGAGACACTGAGACCAAGAGATATCGAGGAACC 480
Db      421  ACACAGAGCATCTCTATAGAAATGAGACACTGAGACCAAGAGATATCGAGGAACC 480
Qy      481  TGCTCCAGTGCATCTGCAAGGCAACGCGCGAGAGAGTGAAGTGTGAGAGGACACCT 540
Db      481  TGCTCCAGTGCATCTGCAAGGCAACGCGCGAGAGAGTGAAGTGTGAGAGGACACCT 540
Qy      541  CTGTGCAAGCACATGAGCGGATCTGGCCCTTCAACCGATGTTGCTATGCTGACCTG 600
Db      541  CTGTGCAAGCACATGAGCGGATCTGGCCCTTCAACCGATGTTGCTATGCTGACCTG 600
Qy      601  AGTGGCTGTAGACCGGTCACTGTCAACCAAGCCAAATGTTGTTAGGGTGTGCTGTA 660
Db      601  AGTGGCTGTAGACCGGTCACTGTCAACCAAGCCAAATGTTGTTAGGGTGTGCTGTA 660
Qy      661  CTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATGATCTAACATCAC 720
Db      661  CTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATGATCTAACATCAC 720
Qy      721  GACCTGCTCATGAGGAGAAAGACAGACCAAGGCTTAAAGTCCAAATCAATTTGCTA 780
Db      721  GACCTGCTCATGAGGAGAAAGACAGACCAAGGCTTAAAGTCCAAATCAATTTGCTA 780
Qy      781  CTGATATGTCGCGAGTGTCAATAAATTGAGAAAGCTGACTTAAGCTATTCAG 840
Db      781  CTGATATGTCGCGAGTGTCAATAAATTGAGAAAGCTGACTTAAGCTATTCAG 840
Qy      841  AACATTTGATCGCTAACGTCACAGTAAACGACGACTCTTGAAGTCAATGATTTTCA 900
Db      841  AACATTTGATCGCTAACGTCACAGTAAACGACGACTCTTGAAGTCAATGATTTTCA 900
Qy      901  GGGATGCAACCATTAAGTGTGAACCGCAAGGCTTACTTGTCTGACCAAAAGTGTTCG 960
Db      901  GGGATGCAACCATTAAGTGTGAACCGCAAGGCTTACTTGTCTGACCAAAAGTGTTCG 960
Qy      961  TAACTTGGCCGACCCCAACCTGTCAGGAATTTTGTCAACGCAATGCTGCGGCTTGAC 1020
Db      961  TAACTTGGCCGACCCCAACCTGTCAGGAATTTTGTCAACGCAATGCTGCGGCTTGAC 1020
Qy      1021  CATATTAAGAAAACCAATACAAAACCAAGCGAAATCTGTGATGTTGAATATATCTGTAC 1080
Db      1021  CATATTAAGAAAACCAATACAAAACCAAGCGAAATCTGTGATGTTGAATATATCTGTAC 1080
Qy      1081  AGTTTACTCCCTTAAACCTGTATGAGATTTGACCAAGGCTCAAAAGATATAGCTAT 1140
Db      1081  AGTTTACTCCCTTAAACCTGTATGAGATTTGACCAAGGCTCAAAAGATATAGCTAT 1140
Qy      1141  TGAAGAACCTAGTATGGTGAACATCATCTTCAAGATTAATGCTCAAGCAGACAA 1200
Db      1141  TGAAGAACCTAGTATGGTGAACATCATCTTCAAGATTAATGCTCAAGCAGACAA 1200
Qy      1201  GGAATTTTAAACAAAACCAACCGAGGCTATGAGTTTATGAAACGATCTCCATATGCTCA 1260

```

```

Db      1201  GCATTTTAAACAAAACCAACCGAGCTATAGCATTTATGAAAGTGAATCTCCATATGCTCA 1260
Qy      1261  CTCAATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGATTTACTACCGTG 1320
Db      1261  |||||
Db      1261  CTCAATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGATTTACTACCGTG 1320
Qy      1321  TTAATAATCGGGAACAAGCTTATAGATCAATTAATAAATCTGCTCTGAATGAAGAAATTA 1380
Db      1321  TTAATAATCGGGAACAAGCTTATAGATCAATTAATAAATCTGCTCTGAATGAAGAAATTA 1380
Qy      1381  ACAACACTGACCTGATCTCTGAGAAATATATAGTCTTTAAAAAGGGGAAAAACCGTATG 1440
Db      1381  ACAACACTGACCTGATCTCTGAGAAATATATAGTCTTTAAAAAGGGGAAAAACCGTATG 1440
Qy      1441  ATCCCTTGTATGCGAGTCACTTTGAACCTGTTCACATCAATTAACCTTGATGCTGATACA 1500
Db      1441  ATCCCTTGTATGCGAGTCACTTTGAACCTGTTCACATCAATTAACCTTGATGCTGATACA 1500
Qy      1501  ACGAATTCCTTAAAGATGAGAGCTCTTAAACAGCTAGCGAAAGCTTAACTTGAATCAGAG 1560
Db      1501  ACGAATTCCTTAAAGATGAGAGCTCTTAAACAGCTAGCGAAAGCTTAACTTGAATCAGAG 1560
Qy      1561  ATTTATAGATCTCTGATATAGGCTTAACTTCTTACACAACTCTGATGCTTTTGGTA 1620
Db      1561  ATTTATAGATCTCTGATATAGGCTTAACTTCTTACACAACTCTGATGCTTTTGGTA 1620
Qy      1621  TTATGACTATACCTTAACTGAGAAAGTATGAGATATACGATGACACCAACCGTATCA 1680
Db      1621  TTATGACTATACCTTAACTGAGAAAGTATGAGATATACGATGACACCAACCGTATCA 1680
Qy      1681  TTAACGTTTATATGAGGCAAGGACCCGAGAGAGAAATGCTAGTACCATTTAGCTGGTG 1740
Db      1681  TTAACGTTTATATGAGGCAAGGACCCGAGAGAGAAATGCTAGTACCATTTAGCTGGTG 1740
Qy      1741  GTGCGCAGGCGCAACAGATTGTACCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGA 1800
Db      1741  GTGCGCAGGCGCAACAGATTGTACCCATAGCTGAGAAAGTGTGTTGATCATGCTGCTGGA 1800
Qy      1801  CTTCTATGTCGTGAGAGAAAGTGGGAGAAAGCCCTACCAAGGCTGATGATGATTAATT 1860
Db      1801  CTTCTATGTCGTGAGAGAAAGTGGGAGAAAGCCCTACCAAGGCTGATGATGATTAATT 1860
Qy      1861  GTACTGCTGTCGAGAGGCGAGCGGACATCACTGTGACTTGTGAATATGATCAACG 1920
Db      1861  GTACTGCTGTCGAGAGGCGAGCGGACATCACTGTGACTTGTGAATATGATCAACG 1920
Qy      1921  ATCAGACACACAGAGACATCTATAGAAATGAGACACTGAGACCAAGAGATTAATCAG 1980
Db      1921  ATCAGACACACAGAGACATCTATAGAAATGAGACACTGAGACCAAGAGATTAATCAG 1980
Qy      1981  GAAACTGCTCAGATGATCTGCAACAGGCAACGCGCAGAGAGTGAAGTGTGAGAGGC 2040
Db      1981  GAAACTGCTCAGATGATCTGCAACAGGCAACGCGCAGAGAGTGAAGTGTGAGAGGC 2040
Qy      2041  ACACCTGTGTCAGACCAATGAGCGGATCTGCGCCCTTCAACGATGTTCTGTAG 2096
Db      2041  ACACCTGTGTCAGACCAATGAGCGGATCTGCGCCCTTCAACGATGTTCTGTAG 2096

```

RESULT 2

US-09-940-235-11

; Sequence 11, Application US/09940235

; Publication No. US20030059921A1

; GENERAL INFORMATION:

; APPLICANT: Kumar, Rajesh

; APPLICANT: Sahni, Girish

; APPLICANT: Roy, Chait

; APPLICANT: Rajagopal, Kammarra

; APPLICANT: Nihalani, Deepak

; APPLICANT: Sundaram, Vasudha

; APPLICANT: Yadav, Mahavir

; TITLE OF INVENTION: NOVEL, CLOT-SPECIFIC SHREPROKINASE

; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION

APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kamraa
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 1541
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-09-940-235-9

Query Match 71.4%; Score 1496.2; DB 3; Length 1541;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 588 ATTGCTGACCTGAGGCGTGTGACCCCTCCATCTGTCAACAACAGCCAAATGTTGTT 647
DB 33 AATGCTGCTGCTGAAATGGCTACTAGATCGCTCTTGTAATAACGCGCAATGGTGTGTT 92
QY 648 AGCGTGTGCTGACTGCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 707
DB 93 AGCGTGTGCTGACTGCTGTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGAATC 152
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA 767
DB 153 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA 212
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCATPAACTTTGAGAAAGTGAATTACTA 827
DB 213 AAACCATTTGCTACTGATAGTGGCGGATGTCATPAACTTTGAGAAAGTGAATTACTA 272
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAAGACGACTTATGAGGTC 887
DB 273 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAAAGACGACTTATGAGGTC 332
QY 888 ATTGATTTTGAAGCGATGCAACCATTAATGATGAAAGCGCAAGGCTTAATTTGCTGAC 947
DB 333 ATTGATTTTGAAGCGATGCAACCATTAATGATGAAAGCGCAAGGCTTAATTTGCTGAC 392
QY 948 AAAGATGCTTGGTAACTTTGCGGACCCCAACCTGTCOAAGATTTTGTCTAAGCGGACAT 1007
DB 333 AAAGATGCTTGGTAACTTTGCGGACCCCAACCTGTCOAAGATTTTGTCTAAGCGGACAT 452
QY 1008 GTGCGGTTAGACCATATTAAGAAAAACAATCAAAACCAAGCAAAATCTGTTGATGTC 1067
DB 453 GTGCGGTTAGACCATATTAAGAAAAACAATCAAAACCAAGCAAAATCTGTTGATGTC 512
QY 1068 GAATATCTGATCAGTTTACTCCCTTAAACCTGATGACGATTTGACACAGGCTCAAA 1127
DB 513 GAATATCTGATCAGTTTACTCCCTTAAACCTGATGACGATTTGACACAGGCTCAAA 572
QY 1128 GATCTAAGCTATTTGAAAACTAGCTATCGGTGACCAATCAATCTCAAGAAATTAATA 1187
DB 573 GATCTAAGCTATTTGAAAACTAGCTATCGGTGACCAATCAATCTCAAGAAATTAATA 632
QY 1188 GGTCAAGCAAAAGATTTTAAACAAAAACCAAGCGGCTATACGATTTATGAACGTGAC 1247

DB 633 GCTCAAGCAAAAGCATTTTAAACAAAAACCAAGGCTATACGATTTATGAACGTGAC 692
QY 1248 TCCCTCAATCTGCTCACTGACGAATGACATTTCCGTACGATTTTACCAATGACCAAGAG 1307
DB 693 TCCCTCAATCTGCTCACTGACGAATGACATTTCCGTACGATTTTACCAATGACCAAGAG 752
QY 1308 TTTACTTACCGTGTGTTAAAAATCCGGAAACAGCTTATAGATCAATTAATAAATCTGTCTG 1367
DB 753 TTTACTTACCGTGTGTTAAAAATCCGGAAACAGCTTATAGATCAATTAATAAATCTGTCTG 812
QY 1368 AATGAAAGAAATTAACCAACTGACCTGATCTCTGAGAAATTAATGCTTTAAAAAGGG 1427
DB 813 AATGAAAGAAATTAACCAACTGACCTGATCTCTGAGAAATTAATGCTTTAAAAAGGG 872
QY 1428 GAAAAAGCGTATGATCCCTTTGATCGAGATCACTTGAACCTGTTCACCATTAATAGTT 1487
DB 873 GAAAAAGCGTATGATCCCTTTGATCGAGATCACTTGAACCTGTTCACCATTAATAGTT 932
QY 1488 GATGTCATACCAAGCAATTTGCTAAAAAGTAGAGAGCTCTTAACAGCTAGGCAAGTAAAC 1547
DB 933 GATGTCATACCAAGCAATTTGCTAAAAAGTAGAGAGCTCTTAACAGCTAGGCAAGTAAAC 992
QY 1548 TTGACTTCAAGATTTTATACGATCTCTGATGATAGGCTAAACTTACTTACCAATCTC 1607
DB 993 TTGACTTCAAGATTTTATACGATCTCTGATGATAGGCTAAACTTACTTACCAATCTC 1052
QY 1608 GATGCTTTTGTATTTATGACATATCTTAACTGGAAGAAAGTAGAGATTAATCAGATGAC 1667
DB 1053 GATGCTTTTGTATTTATGACATATCTTAACTGGAAGAAAGTAGAGATTAATCAGATGAC 1112
QY 1668 ACCAAGCTATCAATTAACGTTTATGAGGAGAGGACCCGAAAGAGAAATGCTAGCTAC 1727
DB 1113 ACCAAGCTATCAATTAACGTTTATGAGGAGAGGACCCGAAAGAGAAATGCTAGCTAC 1172
QY 1728 CATTAGCTGTGTGTGCGCAGGCGCAACAGATTGTACCACTAGCTGAGAAATGTTTGTAT 1787
DB 1173 CATTAGCTGTGTGTGCGCAGGCGCAACAGATTGTACCACTAGCTGAGAAATGTTTGTAT 1232
QY 1788 CATGCTGTGGAATCTTCTATGTTGTGAGAGAAAGTGGAGAAAGCCCTTACCAAGGCTG 1847
DB 1233 CATGCTGTGGAATCTTCTATGTTGTGAGAGAAAGTGGAGAAAGCCCTTACCAAGGCTG 1292
QY 1848 ATGATGTTAATTTGATCTTCCCTGGGAGAGGAGGAGGAGCCATCACTTGTGACCTTCA 1907
DB 1293 ATGATGTTAATTTGATCTTCCCTGGGAGAGGAGGAGGAGCCATCACTTGTGACCTTCA 1352
QY 1908 AATAGATGCAAGATCAGACACAAGGACATCTTATGAATTTGAGACACCTGAGAGCAAG 1967
DB 1353 AATAGATGCAAGATCAGACACAAGGACATCTTATGAATTTGAGACACCTGAGAGCAAG 1412
QY 1968 AAGGATTAATGAGAGAAACCTGCTTCAATGATCTGCAACAGGCAACGCGGAGAGAGTGG 2027
DB 1413 AAGGATTAATGAGAGAAACCTGCTTCAATGATCTGCAACAGGCAACGCGGAGAGAGTGG 1472
QY 2028 AAGTGTGAGAGGACACCTCTGTGACACCAATCGAGCGGATTTGGCCCTTCAACCGAT 2087
DB 1473 AAGTGTGAGAGGACACCTCTGTGACACCAATCGAGCGGATTTGGCCCTTCAACCGAT 1532
QY 2088 GTTCGTTAG 2096
DB 1533 GTTCGTTAG 1541

RESULT 4
US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication No. US20030659921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamraa
; APPLICANT: Nihalani, Deepak

QY 648 AGCGTTGCTGTAAGTGGGGAAGATCAAGATAGTCTTAAATTTTGAATC 707
 DB 193 AGCGTTGCTGTAAGTGGGGAAGATCAAGATAGTCTTAAATTTTGAATC 252
 QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGCTTAACTCCAAATCA 767
 DB 253 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGCTTAACTCCAAATCA 312
 QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGTGAATTCTA 827
 DB 313 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGTGAATTCTA 372
 QY 828 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTACAGCACTTCTTGAAGTGC 887
 DB 373 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTACAGCACTTCTTGAAGTGC 432
 QY 888 ATTGATTTTGAAGCGATGCAACCATTAAGTATGAAAGCGCAAGGTCTACTTGTGCTAC 947
 DB 433 ATTGATTTTGAAGCGATGCAACCATTAAGTATGAAAGCGCAAGGTCTACTTGTGCTAC 492
 QY 948 AAAGATGCTTGGGTAACTTTGCCGACCCCACTGTCAGAAATTTTGTCTAAGCGGACAT 1007
 DB 493 AAAGATGCTTGGGTAACTTTGCCGACCCCACTGTCAGAAATTTTGTCTAAGCGGACAT 552
 QY 1008 GTGGCGCTTGAAGCAATATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTC 1067
 DB 553 GTGGCGCTTGAAGCAATATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTC 612
 QY 1068 GAATATCTGTAAGTCTTCTCCCTTAAACCTGATGAGATTTTCAGACCAAGGTCTCAAA 1127
 DB 613 GAATATCTGTAAGTCTTCTCCCTTAAACCTGATGAGATTTTCAGACCAAGGTCTCAAA 672
 QY 1128 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGCTTAACTCCAAATCA 1187
 DB 673 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGCTTAACTCCAAATCA 732
 QY 1188 GCTCAAGCAAGAAATTTTAAACCAACCCGAGCTTACGATTTTGAACCGTAC 1247
 DB 733 GCTCAAGCAAGAAATTTTAAACCAACCCGAGCTTACGATTTTGAACCGTAC 792
 QY 1248 TCCTCAATGCTCATGATGACATATTCCTGATGAGATTTTCAATGATGATCAAG 1307
 DB 793 TCCTCAATGCTCATGATGACATATTCCTGATGAGATTTTCAATGATGATCAAG 852
 QY 1308 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCTGTCTG 1367
 DB 853 TTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCTGTCTG 912
 QY 1368 AATGAAGAAATTAACAACACTGACCTGATCTGAGAAATTTAGCTCTTAAAGAGG 1427
 DB 913 AATGAAGAAATTAACAACACTGACCTGATCTGAGAAATTTAGCTCTTAAAGAGG 972
 QY 1428 GAAAGACGCTATGATCCCTTGAATGACAGCTTGAACCTGTAACCTTAACTGATGAT 1487
 DB 973 GAAAGACGCTATGATCCCTTGAATGACAGCTTGAACCTTAACTGATGAT 1032
 QY 1488 GATGTCATCAAGCAAGATTTGCTAAATAAGTGAAGCACTTAACTGATGATGATGAT 1547
 DB 1033 GATGTCATCAAGCAAGATTTGCTAAATAAGTGAAGCACTTAACTGATGATGATGAT 1092
 QY 1548 TTTAGCTTGAAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1607
 DB 1093 TTTAGCTTGAAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1152
 QY 1608 GATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1667
 DB 1153 GATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
 QY 1668 AACCAACGCTATCAACCTTTATATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1727
 DB 1213 AACCAACGCTATCAACCTTTATATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1272

QY 1728 CATTTAGCTGATGAT 1742
 DB 1273 CATTTAGCTGATGAT 1287
 RESULT 6
 US-09-940-235-1
 ; Sequence 1, Application US/09940235
 ; Publication No. US2003005921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girdish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammaru
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1245
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1242)
 US-09-940-235-1
 Query Match 54.6%; Score 1145.4; DB 3; Length 1245;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 588 ATTGCTGACCTGATGAGTGGCTGCTGACCGCTGATCTGTCAACACAGCGCAATGTTGTT 647
 DB 1 ATTGCTGACCTGATGAGTGGCTGCTGACCGCTGATCTGTCAACACAGCGCAATGTTGTT 60
 QY 648 AGCGTTGCTGTAAGTGGGGAAGATCAAGATAGTCTTAAATTTTGAATC 707
 DB 61 AGCGTTGCTGTAAGTGGGGAAGATCAAGATAGTCTTAAATTTTGAATC 120
 QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGCTTAACTCCAAATCA 767
 DB 121 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGCTTAACTCCAAATCA 180
 QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGTGAATTCTA 827
 DB 181 AAACCATTTGCTACTGATAGTGGCGGATGTCATATAAATTGAGAAAGTGAATTCTA 240
 QY 828 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTACAGCACTTCTTGAAGTGC 887
 DB 241 AAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTACAGCACTTCTTGAAGTGC 300
 QY 888 ATTGATTTTGAAGCGATGCAACCATTAAGTATGAAAGCGCAAGGTCTACTTGTGCTAC 947
 DB 301 ATTGATTTTGAAGCGATGCAACCATTAAGTATGAAAGCGCAAGGTCTACTTGTGCTAC 360
 QY 948 AAAGATGCTTGGGTAACTTTGCCGACCCCACTGTCAGAAATTTTGTCTAAGCGGACAT 1007
 DB 361 AAAGATGCTTGGGTAACTTTGCCGACCCCACTGTCAGAAATTTTGTCTAAGCGGACAT 420
 QY 1008 GTGGCGCTTGAAGCAATATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTC 1067


```

Qy 1488 GATGTCGATACCAACGAATTGCTAAAAAGTAGACAGCTTTAAACAGTACGAAACGTATAC 1547
    |||||
Db 983 GATGTCGATACCAACGAATTGCTAAAAAGTAGACAGCTTTAAACAGTACGAAACGTATAC 1042
    |||||
Qy 1548 TTGAGCTTCAAGATTTATACGATCTCTCGTATAGGCTTAACTACTCTTCAACATCTC 1607
    |||||
Db 1043 TTGAGCTTCAAGATTTATACGATCTCTCGTATAGGCTTAACTACTCTTCAACATCTC 1102
    |||||
Qy 1608 GATGCTTTGGTATTTATGAGCTATACCTTAACTGGAAGTAGAGATATTCACGATAC 1667
    |||||
Db 1103 GATGCTTTGGTATTTATGAGCTATACCTTAACTGGAAGTAGAGATATTCACGATAC 1162
    |||||
Qy 1668 ACCAACCGTATATTAACCGTTTATATGCGCAAGCGAACCCGAGAGAGAAATGCTAGCTAC 1727
    |||||
Db 1163 ACCAACCGTATATTAACCGTTTATATGCGCAAGCGAACCCGAGAGAGAAATGCTAGCTAT 1222
    |||||
Qy 1728 CATTTAGCTGTGTGT 1742
    |||||
Db 1223 CATTTAGCTGTGTAT 1237
    |||||

```

RESULT 8

```

US-10-474-792-657
; Sequence 657, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winer, Laurie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 657
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-10-474-792-657

```

```

Query Match      47.5%; Score 996; DB 8; Length 1323;
Best Local Similarity 91.3%; Pred. No. 1.9e-277;
Matches 1056; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

```

Qy 587 TATGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAGCCAAATTTGGTGT 646
    |||||
Db 78 TATGCTGAGTATGATGAGTGGCTACAGACCGTCCACCTATCAATAACAGCCAGTTAGTGT 137
    |||||
Qy 647 TAGCGTGTGCTGATCTGTTGAGGGGAGCAATCAAGCATTAAGTCTTAAATTTTGAAT 706
    |||||
Db 138 TAGTATGCGCGGTATCGTTGAAGGTACCGATTAAGTTTATTAATTTTGTGAAT 197
    |||||
Qy 707 CGATCTAACATACGACGCTGCTCATGAGAGAAAGAGAGCAAGCTTTAAGTCCAAAATC 766
    |||||
Db 198 CGATCTAACATACGACGCTGCTCATGAGAGAAAGAGAGCAAGGCTTTAAGTCCAAAATC 257
    |||||
Qy 767 AAAACCATTTGCTACTAGTAGTGGCGAGTGTCAATTAACCTTGAAGAAAGCTGACTTACT 826
    |||||
Db 258 AAAACCATTTGCTACTAGTAGTGGCGAGTGTCAATTAACCTTGAAGAAAGCTGACTTACT 317
    |||||
Qy 827 AAAAGCTATTCAGAAACAATTAATGCTTAACGTCACAGTAAAGCAAGCACTATTGAAGT 886
    |||||
Db 318 AAAAGCTATTCAGAAACAATTAATGCTTAACGTCACAGTAAAGCAAGCACTATTGAAGT 377
    |||||
Qy 887 CATTTATTTTGAAGGAGTGAACCACTTAATGATGCAAAACGGAAGTCTTACTTTGCTGA 946
    |||||
Db 378 CATTTATTTTGAAGGAGTGAACCACTTAATGATGCAAAACGGAAGTCTTACTTTGCTGA 437
    |||||
Qy 947 CAAGAATGTTGCGTAACTTGCAGCAACCAACCTGTCCAAAGATTTTGTAGACGAGCA 1006
    |||||
Db 438 CAAGAATGTTGCGTAACTTGCAGCAACCAACCTGTCCAAAGATTTTGTAGACGAGCA 497
    |||||

```

```

Qy 1007 TGTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTAGT 1066
    |||||
Db 498 TGTGCGGTTAGACCATATATAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTAGT 557
    |||||
Qy 1067 GGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGACAGCGCTTCA 1126
    |||||
Db 558 AGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGAGATTTGACAGCGCTTCA 617
    |||||
Qy 1127 AGATCTAAGCTATTTGAAAAACCTAGTATCGGTGACACATCAATCTCAAGATTTACT 1186
    |||||
Db 618 AGATCTAAGCTATTTGAAAAACCTAGTATCGGTGACACATCAATCTCAAGATTTACT 677
    |||||
Qy 1187 AGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCGAGCTATACGATTTATGAAACGTGA 1246
    |||||
Db 678 AGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCGAGCTATACGATTTATGAAACGTGA 737
    |||||
Qy 1247 CTCCTCAATGCTGACCTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGA 1306
    |||||
Db 738 CTCCTCAATGCTGACCTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGA 797
    |||||
Qy 1307 GTTTACTTACCGTGTAAAAATGCGGAACAAGCTTATAGATCAATTAATAAATCTGCT 1366
    |||||
Db 798 GTTTACTTACCGTGTAAAAATGCGGAACAAGCTTATAGATCAATCTTAACAGAGTAT 857
    |||||
Qy 1367 GAATGAAGAATAATAACAACGACCTGATCTGTGAAATATTAAGTCCCTTAAAAAAG 1426
    |||||
Db 858 TAAAGAAAAACAAGAACCACTGATCTGTCTGAAATAATTAATCTCTTAAAAAAG 917
    |||||
Qy 1427 GAAAAAGCCGTATGATCCCTTTGATGCGACGTCACTTGAACCTGTTCAACCAATACGT 1486
    |||||
Db 918 GAAAAAGCCGTATGATCCCTTTGATGCGACGTCACTTGAACCTGTTCAACCAATACGT 977
    |||||
Qy 1487 TGATGTGATACCAACGAATTTGTTAAAAAGTGAAGAGCTTTTAAACAGTACGAACTGAA 1546
    |||||
Db 978 TGATGTGATACCAACGAATTTGTTAAAAAGTGAAGAGCTTTTAAACAGTACGAACTGAA 1037
    |||||
Qy 1547 CTTAAGCTTCAAGATTTTATAGATCTCTGATTAAGGCTTAACTACTCTTCAACATCT 1606
    |||||
Db 1038 CTTAAGCTTCAAGATTTTATAGATCTCTGATTAAGGCTTAACTACTCTTCAACATCT 1097
    |||||
Qy 1607 CGATGCTTTTGTATTTATGAGTACTTAACTGAAAAAGTAGAGATATACAGATGA 1666
    |||||
Db 1098 CGATGCTTTTGTATTTATGAGTACTTAACTGAAAAAGTAGAGATATATACAGATGA 1157
    |||||
Qy 1667 CACCAACCGTATATTAACGTTTATATGAGCAAGCGACCCGAAAGAGAGATGCTAGCTA 1726
    |||||
Db 1158 GAATTAATCGTGTGTTACATTTATATGAGCAAGCGCCCTTAAAGGGCAAGGCTAGCTA 1217
    |||||
Qy 1727 CCATTTAGCTGTGTGT 1742
    |||||
Db 1218 TCATTTAGCTTTATGAT 1233
    |||||

```

RESULT 9

```

US-09-940-235-3
; Sequence 3, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Gajish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Vadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09

```

PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 777
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(777)
US-09-940-235-3

Query Match 15.7%; Score 330; DB 3; Length 777;
Best Local Similarity 100.0%; Pred. No. 2.1e-84; Indels 0; Gaps 0;
Matches 330; Conservative 0; Mismatches 0;

Qy 1764 CCCATAGCTGAGAAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTGCGAGAAACG 1823
Db 448 CCCATAGCTGAGAAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTGCGAGAAACG 507
Qy 1824 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGCTGATTTGTAATCTTGCTGGAGAAAGGACG 1883
Db 508 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGCTGATTTGTAATCTTGCTGGAGAAAGGACG 567
Qy 1884 GGACCATCTGCACTTCTGAATATGATGCAAGATGAGACAAAGACATCTTAT 1943
Db 568 GGACCATCTGCACTTCTGAATATGATGCAAGATGAGACAAAGACATCTTAT 627
Qy 1944 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 2003
Db 628 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 687
Qy 2004 ACAAGCAACGCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 2063
Db 688 ACAAGCAACGCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 747
Qy 2064 AGCGGATCTGGCCCTTCAACCGATGTTTCT 2093
Db 748 AGCGGATCTGGCCCTTCAACCGATGTTTCT 777

RESULT 10
US-10-210-120-49
Sequence 49, Application US/10210120
Publication No. US20030175736A1
GENERAL INFORMATION:
APPLICANT: Chinnaiyan, Arul M.
APPLICANT: Rubin, Mark A.
APPLICANT: Steekumar, Arun
TITLE OF INVENTION: Expression Profile of Prostate Cancer
FILE REFERENCE: UM-07221
CURRENT APPLICATION NUMBER: US/10/210,120
CURRENT FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: US 60/309,581
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/334,468
PRIOR FILING DATE: 2001-11-15
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn version 3.2
SEQ ID NO 49
LENGTH: 2127
TYPE: DNA
ORGANISM: Homo sapiens
US-10-210-120-49

Query Match 15.7%; Score 330; DB 6; Length 2127;
Best Local Similarity 100.0%; Pred. No. 3.9e-84; Indels 0; Gaps 0;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1764 CCCATAGCTGAGAAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTGCGAGAAACG 1823

Db 597 CCCATAGCTGAGAAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTGCGAGAAACG 656
Qy 1824 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGCTGATTTGTAATCTTGCTGGAGAAAGGACG 1883
Db 657 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGCTGATTTGTAATCTTGCTGGAGAAAGGACG 716
Qy 1884 GGACCATCTGCACTTCTGAATATGATGCAAGATGAGACAAAGACATCTTAT 1943
Db 717 GGACCATCTGCACTTCTGAATATGATGCAAGATGAGACAAAGACATCTTAT 776
Qy 1944 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 2003
Db 777 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 836
Qy 2004 ACAAGCAACGCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 2063
Db 837 ACAAGCAACGCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 896
Qy 2064 AGCGGATCTGGCCCTTCAACCGATGTTTCT 2093
Db 897 AGCGGATCTGGCCCTTCAACCGATGTTTCT 926

RESULT 11
US-10-956-157-4288
Sequence 4288, Application US/10956157
Publication No. US20050118625A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Mounts, William
TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
FILE REFERENCE: 031896-043000 (AM 101081)
CURRENT APPLICATION NUMBER: US/10/956,157
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 319805
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4288
LENGTH: 2127
TYPE: DNA
ORGANISM: Homo sapiens
US-10-956-157-4288

Query Match 15.7%; Score 330; DB 9; Length 2127;
Best Local Similarity 100.0%; Pred. No. 3.9e-84; Indels 0; Gaps 0;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1764 CCCATAGCTGAGAAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTGCGAGAAACG 1823
Db 597 CCCATAGCTGAGAAAGTGTGATCATGCTGCTGGGACTTCCTATGTGTGCGAGAAACG 656
Qy 1824 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGCTGATTTGTAATCTTGCTGGAGAAAGGACG 1883
Db 657 TGGGAGAAAGCCCTTACCAAGGCTGGATGATGCTGATTTGTAATCTTGCTGGAGAAAGGACG 716
Qy 1884 GGACCATCTGCACTTCTGAATATGATGCAAGATGAGACAAAGACATCTTAT 1943
Db 717 GGACCATCTGCACTTCTGAATATGATGCAAGATGAGACAAAGACATCTTAT 776
Qy 1944 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 2003
Db 777 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 836
Qy 2004 ACAAGCAACGCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 2063
Db 837 ACAAGCAACGCGCCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 896
Qy 2064 AGCGGATCTGGCCCTTCAACCGATGTTTCT 2093
Db 897 AGCGGATCTGGCCCTTCAACCGATGTTTCT 926

RESULT 15
US-10-741-601-75
; Sequence 75, Application US/10741601
; Publication No. US2004016519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-75

Query Match 15.7%; Score 330; DB 7; Length 2488;
Best Local Similarity 100.0%; Pred. No. 4.3e-84;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1764 CCATAGCTGAGAGTGTGTTGATCATGCTGCGAATTCTATGTGTGGAAGAACG 1823
|||
DB 913 CCATAGCTGAGAGTGTGTTGATCATGCTGCGAATTCTATGTGTGGAAGAACG 972
|||
QY 1824 TGGAGAGAGCCCTACCAAGCTGATGATGTAGATTGTACTTGCTGCGAGAGGACG 1883
|||
DB 973 TGGAGAGAGCCCTACCAAGCTGATGATGTAGATTGTACTTGCTGCGAGAGGACG 1032
|||
QY 1884 GGACGATCACTTGACCTTCTAGAAATAGATGCAACGATCAGACCAAGACATCCTAT 1943
|||
DB 1033 GGACGATCACTTGACCTTCTAGAAATAGATGCAACGATCAGACCAAGACATCCTAT 1092
|||
QY 1944 AGAATTGAGACACCTGAGCAAGAGATATCGAGAACTGTCTCCAGTGCATCTGC 2003
|||
DB 1093 AGAATTGAGACACCTGAGCAAGAGATATCGAGAACTGTCTCCAGTGCATCTGC 1152
|||
QY 2004 ACAAGCAAGCCGAGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGCAACCAATCG 2063
|||
DB 1153 ACAAGCAAGCCGAGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGCAACCAATCG 1212
|||
QY 2064 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 2093
|||
DB 1213 AGCGGATCTGGCCCTTCAACCGATGTTTGGT 1242
|||

Search completed: February 1, 2006, 14:21:22
Job time : 1732.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 18:17:59 ; Search time 387.523 Seconds
(without alignments)
4492.841 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 2096

Sequence: 1 cgaagacattcatgtgtgtt.....ccttcacgagtcgttag 2096

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6059551 seqs, 415333918 residues

Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2096	100.0	2096	US-10-631-558-12	Sequence 12, App1
2	1684	80.3	1782	US-10-631-558-11	Sequence 11, App1
3	1496.2	71.4	1541	US-10-631-558-9	Sequence 9, App1
4	1165.8	55.6	1661	US-10-631-558-10	Sequence 10, App1
5	1147	54.7	1377	US-10-631-558-5	Sequence 5, App1
6	1145.4	54.6	1245	US-10-631-558-1	Sequence 1, App1
7	1134.2	54.1	1327	US-10-631-558-6	Sequence 6, App1
8	330	15.7	777	US-10-631-558-3	Sequence 3, App1
9	330	15.7	2443	US-10-995-561-114	Sequence 114, App
10	330	15.7	2488	US-10-995-561-105	Sequence 105, App
11	330	15.7	6510	US-10-995-561-112	Sequence 112, App
12	330	15.7	7823	US-10-995-561-117	Sequence 117, App
13	330	15.7	7848	US-10-995-561-111	Sequence 111, App
14	330	15.7	7935	US-10-995-561-113	Sequence 113, App
15	330	15.7	7959	US-10-995-561-108	Sequence 108, App
16	330	15.7	8013	US-10-995-561-104	Sequence 104, App
17	330	15.7	8155	US-10-995-561-116	Sequence 116, App
18	330	15.7	8226	US-10-995-561-107	Sequence 107, App
19	330	15.7	8232	US-10-821-234-693	Sequence 693, App
20	330	15.7	8238	US-10-995-561-106	Sequence 106, App
21	330	15.7	8332	US-10-995-561-110	Sequence 110, App
22	330	15.7	8371	US-10-995-561-109	Sequence 109, App

23	225.8	10.8	8404	US-11-136-527-2446	Sequence 2446, App
C 24	184	8.8	3122	US-11-009-840A-416	Sequence 416, App
C 25	184	8.8	3122	US-11-009-873A-416	Sequence 416, App
C 26	184	8.8	3618	US-11-009-840A-72	Sequence 72, App1
C 27	184	8.8	3618	US-11-009-873A-72	Sequence 72, App1
C 28	184	8.8	4458	US-10-742-634-14	Sequence 14, App1
C 29	184	8.8	4580	US-10-742-634-15	Sequence 15, App1
C 30	184	8.8	5678	US-11-106-820-17	Sequence 17, App1
C 31	184	8.8	5678	US-11-106-820-18	Sequence 18, App1
C 32	184	8.8	5679	US-11-106-820-13	Sequence 13, App1
C 33	184	8.8	5679	US-11-106-820-14	Sequence 14, App1
C 34	184	8.8	12391	US-11-017-550-67	Sequence 67, App1
C 35	184	8.8	16360	US-10-519-531-1	Sequence 1, App1
C 36	184	8.8	17207	US-10-519-531-8	Sequence 8, App1
C 37	182.4	8.7	11627	US-11-220-995-1	Sequence 1, App1
C 38	181	8.6	9808	US-10-966-483-43	Sequence 43, App1
C 39	181	8.6	9808	US-11-021-441-33	Sequence 33, App1
C 40	160.8	7.7	87672	US-10-995-561-13237	Sequence 13237, A
C 41	160.6	7.7	5347	US-11-203-589-1	Sequence 1, App1
C 42	160.6	7.7	8031	US-10-623-155-254	Sequence 254, App
C 43	100.4	4.8	4974	US-11-094-586-17	Sequence 17, App1
C 44	100.4	4.8	4974	US-11-076-733-86	Sequence 86, App1
C 45	100.4	4.8	8966	US-11-076-733-87	Sequence 87, App1

ALIGNMENTS

RESULT 1
US-10-631-558-12
Sequence 12, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammarra
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
FILE REFERENCE: 07064-00902
CURRENT APPLICATION NUMBER: US/10/631,558
CURRENT FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 2096
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-10-631-558-12

Query Match 100.0%; Score 2096; DB 7; Length 2096;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 2096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CGAAGACATTTCATGTTGTCAGGTCGAGACGTTTTCACAGCAGTCGCTTACG 60
Db 1 CGAAGACATTTCATGTTGTCAGGTCGAGACGTTTTCACAGCAGTCGCTTACG 60
Cy 61 TTGCTGCGGTATCGGTGATTCATTGCTGCTTACAGGTAAGCAACCCGCGACCTTACC 120

Db 61 TTCGCTCGGCTATCGGTATGATTAATCTGCTAACAGTAAGGCAACCCCGCAGCTAGCC 120
 Qy 121 GGGTCTCAACGACAGAGACAGATCATGCGACCCGTGGCCAGAACCCAAACGCTGCCG 180
 Db 121 GGGTCTCAACGACAGAGACAGATCATGCGACCCGTGGCCAGAACCCAAACGCTGCCG 180
 Qy 181 AGATCTCGATCCCGGAAATTAATACGACTCACTATAGGGAACAACGCTTTCCCTC 240
 Db 181 AGATCTCGATCCCGGAAATTAATACGACTCACTATAGGGAACAACGCTTTCCCTC 240
 Qy 241 TAGAAATATTTTGTATTAATTAAAGAGAGATATACCATGCTGGAAGACAAACAGATT 300
 Db 241 TAGAAATATTTTGTATTAATTAAAGAGAGATATACCATGCTGGAAGACAAACAGATT 300
 Qy 301 GTACCACTAGCTGAGAAAGTGTATGATCATGCTGCTGGACTTCTATGCTGCGAGAA 360
 Db 301 GTACCACTAGCTGAGAAAGTGTATGATCATGCTGCTGGACTTCTATGCTGCGAGAA 360
 Qy 361 ACCTGGAG 420
 Db 361 ACCTGGAG 420
 Qy 421 ACAG 480
 Db 421 ACAG 480
 Qy 481 TGCTCAAGTGCATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 Db 481 TGCTCAAGTGCATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 Qy 541 CTGTGACAG 600
 Db 541 CTGTGACAG 600
 Qy 601 AGTGGCTGTAG 660
 Db 601 AGTGGCTGTAG 660
 Qy 661 CTGTGAG 720
 Db 661 CTGTGAG 720
 Qy 721 GACCTGCTATGAG 780
 Db 721 GACCTGCTATGAG 780
 Qy 781 CTGATAGTGGCGAGATGTCACATTAACCTTGAAGAGCTGACTTAAGAGCTATTCAG 840
 Db 781 CTGATAGTGGCGAGATGTCACATTAACCTTGAAGAGCTGACTTAAGAGCTATTCAG 840
 Qy 841 AACATTTGATGCTTAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Db 841 AACATTTGATGCTTAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Qy 901 GCGATGCAACATTAATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 Db 901 GCGATGCAACATTAATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 Qy 961 TAACTTGGCGAG 1020
 Db 961 TAACTTGGCGAG 1020
 Qy 1021 CATATTAAG 1080
 Db 1021 CATATTAAG 1080
 Qy 1081 AGTTACTCTCTTAAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 Db 1081 AGTTACTCTCTTAAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 Qy 1141 TGAAG 1200
 Db 1141 TGAAG 1200

Qy 1201 GCATTTTAAACAAAACCAACCCAGCTATACGATTTATGAACTGACTCTCAATGCTCA 1260
 Db 1201 GCATTTTAAACAAAACCAACCCAGCTATACGATTTATGAACTGACTCTCAATGCTCA 1260
 Qy 1261 CTATGACATGACATTTTCCGTACGATTTTACGATGATGACAGAGATTTTACCTTAC 1320
 Db 1261 CTATGACATGACATTTTCCGTACGATTTTACGATGATGACAGAGATTTTACCTTAC 1320
 Qy 1321 TTAAG 1380
 Db 1321 TTAAG 1380
 Qy 1381 ACAACAG 1440
 Db 1381 ACAACAG 1440
 Qy 1441 ATCCCTTTGATCGAGTCACTTGAACCTGTTACATCAATCAATGATGATGATGATGAT 1500
 Db 1441 ATCCCTTTGATCGAGTCACTTGAACCTGTTACATCAATCAATGATGATGATGATGAT 1500
 Qy 1501 ACGAATGCTTAAAG 1560
 Db 1501 ACGAATGCTTAAAG 1560
 Qy 1561 ATTTATGAGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Db 1561 ATTTATGAGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Qy 1621 TTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 1621 TTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Qy 1681 TTAACGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
 Db 1681 TTAACGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
 Qy 1741 GTGGCAG 1800
 Db 1741 GTGGCAG 1800
 Qy 1801 CTTCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 Db 1801 CTTCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 Qy 1861 GTACTTCTGAG 1920
 Db 1861 GTACTTCTGAG 1920
 Qy 1921 ATCAG 1980
 Db 1921 ATCAG 1980
 Qy 1981 GAAACCTGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Db 1981 GAAACCTGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Qy 2041 ACACCTTGTGAG 2096
 Db 2041 ACACCTTGTGAG 2096

RESULT 2
 US-10-631-558-11
 ; Sequence 11, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sami, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammarra
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha

APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
CURRENT FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1782
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-10-631-558-11

Query Match 80.3%; Score 1684; DB 7; Length 1782;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 51 TCGCTTACGTTCCGCTCGGCTATCGGTGATTCATTCTGCTAACCAAGTAAGCAACCCCGC 110
DB 1 TCGCTTACGTTCCGCTCGGCTATCGGTGATTCATTCTGCTAACCAAGTAAGCAACCCCGC 60
QY 111 CAGCTAGCCGGGCTCTCAACGACAGAGACAGATCATGCGGACCCGTCGCGACGACCA 170
DB 61 CAGCTAGCCGGGCTCTCAACGACAGAGACAGATCATGCGGACCCGTCGCGACGACCA 120
QY 171 ACGCTGCCGAGATCTCGATCCCGCGAAATTAATGAGCTCATTAAGGAGACCAAC 230
DB 121 ACGCTGCCGAGATCTCGATCCCGCGAAATTAATGAGCTCATTAAGGAGACCAAC 180
QY 231 GGTTCCTCTAGAAATTAATTTTGTATTAATTAAAGAGATATACATGCTGCAAGC 290
DB 181 GGTTCCTCTAGAAATTAATTTTGTATTAATTAAAGAGATATACATGCTGCAAGC 240
QY 291 ACAAGATTTGATCCCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCTATGT 350
DB 241 ACAAGATTTGATCCCATAGCTGAGAGAGTTTGTATCATGCTGCGGACCTTCTATGT 300
QY 351 GGTGAGAGAAAGTGGAGAGAGGACGACGATCATCTTGCACTTCTAGAAATTAAGTGC 410
DB 301 GGTGAGAGAAAGTGGAGAGAGGACGACGATCATCTTGCACTTCTAGAAATTAAGTGC 360
QY 411 AACGATCAGACACAAAGACATCTTATAGAAATTTGAGACACCTGAGCAAGAGATTAAT 470
DB 361 AACGATCAGACACAAAGACATCTTATAGAAATTTGAGACACCTGAGCAAGAGATTAAT 420
QY 471 CGAGGAAACCTGCTCAAGTGCATCTGCAACGACGCGCGAGAGAGTGAAGTGTAG 530
DB 421 CGAGGAAACCTGCTCAAGTGCATCTGCAACGACGCGCGAGAGAGTGAAGTGTAG 480
QY 531 AGGCAACACTCTGTGTCACACCATGTCGAGACCGATCTGGCCCTTCAACCGATTTCCGTAAT 590
DB 481 AGGCAACACTCTGTGTCACACCATGTCGAGACCGATCTGGCCCTTCAACCGATTTCCGTAAT 540
QY 591 GCTGACCTGAGTGTGCTGTAGACGCTCATGTCATGTCACAAACGACCAATTTGTTAGC 650
DB 541 GCTGACCTGAGTGTGCTGTAGACGCTCATGTCATGTCACAAACGACCAATTTGTTAGC 600
QY 651 GTTGTGCTACTGTGAGGAGGAGCAATCAAGACATTAGTCTTAAATTTTGTGAATCGAT 710
DB 601 GTTGTGCTACTGTGAGGAGGAGCAATCAAGACATTAGTCTTAAATTTTGTGAATCGAT 660
QY 711 CTAACATCAACGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCAAAA 770

DB 661 CTAACATCAACGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCAAAA 720
QY 771 CCAATTTGCTATGTAAGTGGGCGCATGTCACATTAACCTTGAGAAAGCTGACTTAAAG 830
DB 721 CCAATTTGCTATGTAAGTGGGCGCATGTCACATTAACCTTGAGAAAGCTGACTTAAAG 780
QY 831 GCTATTCAGAAACATTAAGTGTGCTTACGTCACAGTAACGACCTACTTGAAGTCAAT 890
DB 781 GCTATTCAGAAACATTAAGTGTGCTTACGTCACAGTAACGACCTACTTGAAGTCAAT 840
QY 891 GATTTTCAGACGATGCAACCATTAAGTATGAAAACGCGCAAGTCTACTTTCGTCGACAA 950
DB 841 GATTTTCAGACGATGCAACCATTAAGTATGAAAACGCGCAAGTCTACTTTCGTCGACAA 900
QY 951 GATGCTTCGTAACCTTGGCCGACCCCACTCTCCAAAGATTTTGTCTAAGCGACATGTG 1010
DB 901 GATGCTTCGTAACCTTGGCCGACCCCACTCTCCAAAGATTTTGTCTAAGCGACATGTG 960
QY 1011 CGGTTAGACCATTAAGAAAGAAACCAATACAAACCAAGGGAATCTGTAGTGTGAA 1070
DB 961 CGGTTAGACCATTAAGAAAGAAACCAATACAAACCAAGGGAATCTGTAGTGTGAA 1020
QY 1071 TATACGTACAGTTTACTCTCTTAAACCTGATGACGATTTCAACACAGTCTCAAGAT 1130
DB 1021 TATACGTACAGTTTACTCTCTTAAACCTGATGACGATTTCAACACAGTCTCAAGAT 1080
QY 1131 ACTAAGCTATTGAAAACATAGCTATGCTGTCACACCATCACTCAAGAAATTAAGTCT 1190
DB 1081 ACTAAGCTATTGAAAACATAGCTATGCTGTCACACCATCACTCAAGAAATTAAGTCT 1140
QY 1191 CAAGCAAGACATTTTAAACAAACCAACCAAGCTATAGATTTAAGACGTGATCC 1250
DB 1141 CAAGCAAGACATTTTAAACAAACCAACCAAGCTATAGATTTAAGACGTGATCC 1200
QY 1251 TCAATGCTCACTGACATGACATTTTCCGTGATTTTAAACCAAGTCAAGAGTTT 1310
DB 1201 TCAATGCTCACTGACATGACATTTTCCGTGATTTTAAACCAAGTCAAGAGTTT 1260
QY 1311 ACTTACGTTTAAAAATCGGGAACAGCTTATAGATCAATTAATAATCTGCTGAT 1370
DB 1261 ACTTACGTTTAAAAATCGGGAACAGCTTATAGATCAATTAATAATCTGCTGAT 1320
QY 1371 GAAGAAATTAACCAACATGACCTGATCTGAGAAATTTAGTCTTAAAAAGGGGA 1430
DB 1321 GAAGAAATTAACCAACATGACCTGATCTGAGAAATTTAGTCTTAAAAAGGGGA 1380
QY 1431 AAGCCGATGATCCCTTGAATGCGAGTCACTTGAACCTGTCACATCAATTAAGTGTAT 1490
DB 1381 AAGCCGATGATCCCTTGAATGCGAGTCACTTGAACCTGTCACATCAATTAAGTGTAT 1440
QY 1491 GTGATATCAACGAATTTGCTAAAGAGTACAGCTCTTAACAGCTACGAACTTAAGT 1550
DB 1441 GTGATATCAACGAATTTGCTAAAGAGTACAGCTCTTAACAGCTACGAACTTAAGT 1500
QY 1551 GACTTCAGAGATTTATAGATCTCTGATTAAGGCTTAACTACTTCAACCAATCTGAT 1610
DB 1501 GACTTCAGAGATTTATAGATCTCTGATTAAGGCTTAACTACTTCAACCAATCTGAT 1560
QY 1611 GCTTTTGTATTTATAGATCTTATACCTTAACTGGAAGATTAATCAAGATGACACC 1670
DB 1561 GCTTTTGTATTTATAGATCTTATACCTTAACTGGAAGATTAATCAAGATGACACC 1620
QY 1671 AACCGTATCAATACCGTTTATATAGGGAAGACCGAAGGAAGAAATGCTAGTACAT 1730
DB 1621 AACCGTATCAATACCGTTTATATAGGGAAGACCGAAGGAAGAAATGCTAGTACAT 1680
QY 1731 TTAGCTGTGCT 1742
DB 1681 TTAGCTGTGCT 1692

RESULT 3

US-10-631-558-9
; Sequence 9, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumari, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-9

Query Match 71.4%; Score 1496.2; DB 7; Length 1541;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 588 ATTGCTGACCTGAGTGGTGTGCTAGACCGTCCATCTGTCAACAACGCCAATTGGTTGT 647
DB 33 ATAGCTGCTCCTGGAATGCTACTAGATCGTCTCTGTAATAACGCAATGGTGTGT 92
QY 648 AGCGTGTGCTGACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTTTGAATC 707
DB 93 AGCGTGTGCTGACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTTTGAATC 152
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA 767
DB 153 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGGCTTAAGTCCAAATCA 212
QY 768 AAACCATTTGCTACTGATATGTCGCGGATGTCAATTAACCTTGAGAAAGCTGACTTA 827
DB 213 AAACCATTTGCTACTGATATGTCGCGGATGTCAATTAACCTTGAGAAAGCTGACTTA 272
QY 828 AAGGCTATTCAAGAACAAATGATGCTAAGTCCACAGTAAGACGACTATTGAGGTC 887
DB 273 AAGGCTATTCAAGAACAAATGATGCTAAGTCCACAGTAAGACGACTATTGAGGTC 332
QY 888 ATTGATTTTGGCAAGGATGCAACCATTAAGTCAAGACGAGGCTTAAGTCTGAGC 947
DB 333 ATTGATTTTGGCAAGGATGCAACCATTAAGTCAAGACGAGGCTTAAGTCTGAGC 392
QY 948 AAAGATGTTGGTGAACCTTGCAGACCAACCTGTCAGAAATTTTGGCTAAGCGGACAT 1007
DB 393 AAAGATGTTGGTGAACCTTGCAGACCAACCTGTCAGAAATTTTGGCTAAGCGGACAT 452
QY 1008 GTGCGCGTTAGCAATATTAAGAAAAACCAATACAAACCAAGCGCAATCTGTGATGTC 1067
DB 453 GTGCGCGTTAGCAATATTAAGAAAAACCAATACAAACCAAGCGCAATCTGTGATGTC 512
QY 1068 GAATATACGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1127
DB 513 GAATATACGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 572

QY 1128 GATCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1187
DB 573 GATCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 632
QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAAMACACCCAGGCTATACGATTTTGAACGTGAC 1247
DB 633 GCTCAAGCAAAAGCAATTTTAAACAAAMACACCCAGGCTATACGATTTTGAACGTGAC 692
QY 1248 TCTCTCAATGCTCACTGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 693 TCTCTCAATGCTCACTGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 752
QY 1308 TTACTCAACGCTGTTAAAAATCGGGAAACAGCTTATGATCAATAAAAAATCGTCTG 1367
DB 753 TTACTCAACGCTGTTAAAAATCGGGAAACAGCTTATGATCAATAAAAAATCGTCTG 812
QY 1368 AATGAAGAAATTAACAACATGACCTGATCTGAGAAATTAACGTCCTTAAAAAAGG 1427
DB 813 AATGAAGAAATTAACAACATGACCTGATCTGAGAAATTAACGTCCTTAAAAAAGG 872
QY 1428 GAAAAAGCGTATGATCCCTTTGATTCGACATCACTTGAACCTTTACCATCAATACCT 1487
DB 873 GAAAAAGCGTATGATCCCTTTGATTCGACATCACTTGAACCTTTACCATCAATACCT 932
QY 1488 GATGTGATACCAACGAAATGCTTAAAAAGTACGAGCTTAAACAGCTAGGAAACGTAC 1547
DB 933 GATGTGATACCAACGAAATGCTTAAAAAGTACGAGCTTAAACAGCTAGGAAACGTAC 992
QY 1548 TTAGACTTCAGAGATTTATACATCTCGATCTGATAGGCTTAACTTCAACATCTC 1607
DB 993 TTAGACTTCAGAGATTTATACATCTCGATCTGATAGGCTTAACTTCAACATCTC 1052
QY 1608 GATGCTTTGGTATTTATGACATTAACCTTAACCTGAAAAAGTAAAGATTAACGATGAC 1667
DB 1053 GATGCTTTGGTATTTATGACATTAACCTTAACCTGAAAAAGTAAAGATTAACGATGAC 1112
QY 1668 ACCAACCGTATCAATACCGTTTATATGAGGACGACCCGAAAGAGAAATGCTAGTAC 1727
DB 1113 ACCAACCGTATCAATACCGTTTATATGAGGACGACCCGAAAGAGAAATGCTAGTAC 1172
QY 1728 CATTTAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1787
DB 1173 CATTTAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1232
QY 1788 CATGCTGTGGAACCTTCTATGTCGAGAAACGTGGAGAAAGCCCTTACCAAGGCTG 1847
DB 1233 CATGCTGTGGAACCTTCTATGTCGAGAAACGTGGAGAAAGCCCTTACCAAGGCTG 1292
QY 1848 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1907
DB 1293 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1352
QY 1908 AATAGATGCAACGATCAGGACCAAGGACATCTTATGATTTGAGACACTGAGGAA 1967
DB 1353 AATAGATGCAACGATCAGGACCAAGGACATCTTATGATTTGAGACACTGAGGAA 1412
QY 1968 AAGGATTAATGAGAAACCTGCTCAAGTGCATCTGCAAGGCAACGCGCGAGAGATG 2027
DB 1413 AAGGATTAATGAGAAACCTGCTCAAGTGCATCTGCAAGGCAACGCGCGAGAGATG 1472
QY 2028 AAGTGAAGGACCACTTGTGCAAGACCAATCGAGCGGATGTCGCTTCAACGAT 2087
DB 1473 AAGTGAAGGACCACTTGTGCAAGACCAATCGAGCGGATGTCGCTTCAACGAT 1532
QY 2088 GTTGGTTAG 2096
DB 1533 GTTGGTTAG 1541

RESULT 4
US-10-631-558-10
; Sequence 10, Application US/10631558

Publication No.: US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammar
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEIN POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
PRIOR FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 1661
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-10-631-558-10

Query Match 55.6%; Score 1165.8; DB 7; Length 1661;
Beet Local Similarity 99.4%; Pred. No. 0;
Matches 1170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 588 ATTGCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCAACAACAGCCCAATTGGTGT 647
DB 183 ATTGCTGACCTGAGTGGCTGCTGACCGTCCATCTGTCAACAACAGCCCAATTGGTGT 242
QY 648 AGCGTTGCTGTACTGTGAGGGGACGATCAAGACATTAGTCTTAAATTTTGAATC 707
DB 243 AGCGTTGCTGTACTGTGAGGGGACGATCAAGACATTAGTCTTAAATTTTGAATC 302
QY 708 GATCTAACATCAAGACCTGCTCATGAGAAACAGACAGAGCTTAAGTCCAAATCA 767
DB 303 GATCTAACATCAAGACCTGCTCATGAGAAACAGACAGAGCTTAAGTCCAAATCA 362
QY 768 AAACATTTGCTACTAGTAGTGGCGGATGTCACTAACTTGAGAAAGCTGACTTA 827
DB 363 AAACATTTGCTACTAGTAGTGGCGGATGTCACTAACTTGAGAAAGCTGACTTA 422
QY 828 AAGGCTATTCAAGAACATTTGATGCTAAAGTCCACAGTACGACGACTTCTTGAAGTC 887
DB 423 AAGGCTATTCAAGAACATTTGATGCTAAAGTCCACAGTACGACGACTTCTTGAAGTC 482
QY 888 ATTGATTTTGAAGGACGATGACCACTTACTGATGAAACGGCAAGTCTTACTTGTGAC 947
DB 483 ATTGATTTTGAAGGACGATGACCACTTACTGATGAAACGGCAAGTCTTACTTGTGAC 542
QY 948 AAAGATGTTGGGTAACTTGGCGACCCAACTGCTCAAGAAATTTTGTCAACGGACAT 1007
DB 543 AAAGATGTTGGGTAACTTGGCGACCCAACTGCTCAAGAAATTTTGTCAACGGACAT 602
QY 1008 GTGGCGGTGACCATATAAAGAAAAACCAATCAAAACCAAGCAAGAAATCTGTGATGTG 1067
DB 603 GTGGCGGTGACCATATAAAGAAAAACCAATCAAAACCAAGCAAGAAATCTGTGATGTG 662
QY 1068 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACAGGTCTCAAA 1127
DB 663 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACAGGTCTCAAA 722
QY 1128 GATATCAAGCTATGAAACACATAGCTATGCTGAGACCAATCAATCTCAAGAAATTA 1187

DB 723 GATATCAAGCTATGAAACACATAGCTATGCTGAGACCAATCAATCTCAAGAAATTA 782
QY 1188 GCTAAGACAAAGCATTTTAAACAAAACCAACCGGCTATGACATTTATGACGTAC 1247
DB 783 GCTAAGACAAAGCATTTTAAACAAAACCAACCGGCTATGACATTTATGACGTAC 842
QY 1248 TCTCAATGCTACCTGATGACAAATGATTTCCGTAGATTTTCAATGATGACAG 1307
DB 843 TCTCAATGCTACCTGATGACAAATGATTTCCGTAGATTTTCAATGATGACAG 902
QY 1308 TTACTTACCGTGTAAATTCGGACAAAGCTTATGATCAATAAATTTCTGTCTG 1367
DB 903 TTACTTACCGTGTAAATTCGGACAAAGCTTATGATCAATAAATTTCTGTCTG 962
QY 1368 AATGAAGAAATTAACACACTGACTGATCTCTGAAATATTTAGTCTTAAAAAGG 1427
DB 963 AATGAAGAAATTAACACACTGACTGATCTCTGAAATATTTAGTCTTAAAAAGG 1022
QY 1428 GAAAGCCGTATGATCCCTTATGTCAGTCACTTGAACATGATCAATGATGAT 1487
DB 1023 GAAAGCCGTATGATCCCTTATGTCAGTCACTTGAACATGATCAATGATGAT 1082
QY 1488 GATGCTATCAACCAAGATTTCTAAATAGTACAGCTCTTAAACAGTACGAA 1547
DB 1083 GATGCTATCAACCAAGATTTCTAAATAGTACAGCTCTTAAACAGTACGAA 1142
QY 1548 TTAGACTTCAGAGATTTATGATCTCTGTATTAAGCTTAACCTTCAACATCTC 1607
DB 1143 TTAGACTTCAGAGATTTATGATCTCTGTATTAAGCTTAACCTTCAACATCTC 1202
QY 1608 GATGCTTTGCTATTAATGACTTATACCTTAATGAAAGTATGAGATTAACAGATAC 1667
DB 1203 GATGCTTTGCTATTAATGACTTATACCTTAATGAAAGTATGAGATTAACAGATAC 1262
QY 1668 ACCAAGCTATCATTAACCGTTTATATGAGGACGACCCGAAAGGAGAAATGCTAGTAC 1727
DB 1263 ACCAAGCTATCATTAACCGTTTATATGAGGACGACCCGAAAGGAGAAATGCTAGTAC 1322
QY 1728 CATTTAGCCGTGTGTCGACGAGCGCAACGATTTGAC 1764
DB 1323 CATTTAGCCGTGTGTCGACGAGCGCAACGATTTGAC 1359
RESULT 5
US-10-631-558-5
Sequence 5, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammar
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEIN POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
PRIOR FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US/09/940,235
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1377

TYPE: DNA
ORGANISM: Streptococcus equisimilis
US-10-631-558-5

Query Match 54.7%; Score 1147; DB 7; Length 1377;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 588 ATTGCTGAGCGTGAAGGCGCTGAGACCGTCCATCTGTGACAAACAGCCAAATGGTGTG 647
DB 133 ATTGCTGAGCGTGAAGGCGCTGAGACCGTCCATCTGTGACAAACAGCCAAATGGTGTG 192
QY 648 AGCGTTGCTGTAAGTGTGAGGGAAGCAATCAAGACTTGTCTTAATTTTGAATC 707
DB 193 AGCGTTGCTGTAAGTGTGAGGGAAGCAATCAAGACTTGTCTTAATTTTGAATC 252
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
DB 253 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 312
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTA 827
DB 313 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTA 372
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAAAGTCCACAGTAAAGCACTACTTGAAGTC 887
DB 373 AAGGCTATTCAAGAACATTTGATCGCTAAAGTCCACAGTAAAGCACTACTTGAAGTC 432
QY 888 ATTGATTTTGAAGCGAGTCAACCATTTACTGATGAAACGGCAAGGTCTACTTGTGAC 947
DB 433 ATTGATTTTGAAGCGAGTCAACCATTTACTGATGAAACGGCAAGGTCTACTTGTGAC 492
QY 948 AAAGATGTTGCGGTAACTTGGCCGACCCACCTGTCCAGAAATTTTGTAGCGGACAT 1007
DB 493 AAAGATGTTGCGGTAACTTGGCCGACCCACCTGTCCAGAAATTTTGTAGCGGACAT 552
QY 1008 GTGCGCGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTC 1067
DB 553 GTGCGCGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTC 612
QY 1068 GAATTAATCTGTAAGTGTGATGTCCTCTTAAACCTGATGACGATTTTGACAGCAGTCTCAA 1127
DB 613 GAATTAATCTGTAAGTGTGATGTCCTCTTAAACCTGATGACGATTTTGACAGCAGTCTCAA 672
QY 1128 GATTAATGCTATTTGAAACCACTAGTATGCGTGACACATCACTCTCAAGATTA 1187
DB 673 GATTAATGCTATTTGAAACCACTAGTATGCGTGACACATCACTCTCAAGATTA 732
QY 1188 GCTCAAGCAAGGATTTTAAACAAACCAACCGGCTATTCGATTTATGAAACGTGAC 1247
DB 733 GCTCAAGCAAGGATTTTAAACAAACCAACCGGCTATTCGATTTATGAAACGTGAC 792
QY 1248 TCCCTCAATCGTCACTCATGACATGATCATTTTCCGATGATTTTACCAATGATCAAGAG 1307
DB 793 TCCCTCAATCGTCACTCATGACATGATCATTTTCCGATGATTTTACCAATGATCAAGAG 852
QY 1308 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCTGCTTC 1367
DB 853 TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCTGCTTC 912
QY 1368 AATGAAGAAATTAACAACAACCTGATCTTGAGAAATATTTAAGTCTTAAATAAAGG 1427
DB 913 AATGAAGAAATTAACAACAACCTGATCTTGAGAAATATTTAAGTCTTAAATAAAGG 972
QY 1428 GAAAGACGATGATCCCTTGTGACGATCACTTGAACCTGTTCAACATCAATAGCTT 1487
DB 973 GAAAGACGATGATCCCTTGTGACGATCACTTGAACCTGTTCAACATCAATAGCTT 1032
QY 1488 GATGTCGATCAACGAATTTGCTAAAGATGAGCAGCTTTAAACAGCTAGCGAAGCTAAC 1547
DB 1033 GATGTCGATCAACGAATTTGCTAAAGATGAGCAGCTTTAAACAGCTAGCGAAGCTAAC 1092
QY 1548 TTAGACTTCAAGGATTTTATGATCTCGTGAATAGGCTTAAGTCTTCAACAACATCTC 1607

DB 1093 TTAGACTTCAAGGATTTTATGATCTCGTGAATAGGCTTAAGTCTTCAACAACATCTC 1152
QY 1608 GATGCTTTTGGTATTTATGACATTAACCTTAACTGGAAGATGAGATTAACGATGAC 1667
DB 1153 GATGCTTTTGGTATTTATGACATTAACCTTAACTGGAAGATGAGATTAACGATGAC 1212
QY 1668 ACCAAGCTATTCATTAACCGTTTATAGGGAAGGACCCGAAGGAGAAATGCTAGCTAC 1727
DB 1213 ACCAAGCTATTCATTAACCGTTTATAGGGAAGGACCCGAAGGAGAAATGCTAGCTAT 1272
QY 1728 CATTTAGCTGATGAT 1742
DB 1273 CATTTAGCTGATGAT 1287

RESULT 6
US-10-631-558-1
Sequence 1, Application US/10631558
Publication No. US20050260598A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kamara
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/10/631,558
CURRENT FILING DATE: 2003-07-31
PRIORITY APPLICATION NUMBER: US/09/940,235
PRIORITY FILING DATE: 2002-04-09
PRIORITY APPLICATION NUMBER: 09/471,349
PRIORITY FILING DATE: 1999-12-23
PRIORITY APPLICATION NUMBER: IN 3825/DEL/98
PRIORITY FILING DATE: 1998-12-24
NUMBER OF SEQ. ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ. ID NO. 1
LENGTH: 1245
TYPE: DNA
ORGANISM: Streptococcus equisimilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1242)
US-10-631-558-1

Query Match 54.6%; Score 1145.4; DB 7; Length 1245;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 588 ATTGCTGAGCGTGAAGGCGCTGAGACCGTCCATCTGTGACAAACAGCCAAATGGTGTG 647
DB 1 ATTGCTGAGCGTGAAGGCGCTGAGACCGTCCATCTGTGACAAACAGCCAAATGGTGTG 60
QY 648 AGCGTTGCTGTAAGTGTGAGGGAAGCAATCAAGACTTGTCTTAATTTTGAATC 707
DB 61 AGCGTTGCTGTAAGTGTGAGGGAAGCAATCAAGACTTGTCTTAATTTTGAATC 120
QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
DB 121 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 180
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTA 827
DB 181 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTA 240
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAAAGTCCACAGTAAAGCACTACTTGAAGTC 887


```

Db      ||||| 241 AAGGCTATTCAAAACAAATTGATGCTMAAGCTCCACGTAACGACGCTACTTTGAGGTC 300
Oy      ||||| 888 ATTGATTTTGCAGACGATGCAACCAATTACTGATGCAAAAGGCAAGCTACTTTGCGAC 947
Db      ||||| 301 ATTGATTTTGCAGACGATGCAACCAATTACTGATGCAAAAGGCAAGCTACTTTGCGAC 360
Oy      ||||| 948 AAAGATGTTGCGTAACTTCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGACAT 1007
Db      ||||| 361 AAAGATGTTGCGTAACTTCCGACCAACCTGTCCAGAAATTTTGTCTAAGCGACAT 420
Oy      ||||| 1008 GTGGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 1067
Db      ||||| 421 GTGGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 480
Oy      ||||| 1068 GAATATATCTGTACAGTTTCTCCCTTAAACCTGTGACGATTTTCAACAGGCTCAAA 1127
Db      ||||| 481 GAATATATCTGTACAGTTTCTCCCTTAAACCTGTGACGATTTTCAACAGGCTCAAA 540
Oy      ||||| 1128 GATATCTAAGCTATTGAAAAACATAGCTATGCGTGAACCATCACTCAAGAAATTACTA 1187
Db      ||||| 541 GATATCTAAGCTATTGAAAAACATAGCTATGCGTGAACCATCACTCAAGAAATTACTA 600
Oy      ||||| 1188 GCTCAAGCACAAAACATTTTAAACAAAACCAACCAGGCTATACGATTTTGAACGTCAC 1247
Db      ||||| 601 GCTCAAGCACAAAACATTTTAAACAAAACCAACCAGGCTATACGATTTTGAACGTCAC 660
Oy      ||||| 1248 TCCTCAATGCTCACTCATGACATGACATTTTCCGTAACGATTTTCAAAATGATCAAGAG 1307
Db      ||||| 661 TCCTCAATGCTCACTCATGACATGACATTTTCCGTAACGATTTTCAAAATGATCAAGAG 720
Oy      ||||| 1308 TTTACTTACCGTGTAAATAAGGGAAACAAGCTTATAGATCAATTAATAAATCGGTG 1367
Db      ||||| 721 TTTACTTACCGTGTAAATAAGGGAAACAAGCTTATAGATCAATTAATAAATCGGTG 780
Oy      ||||| 1368 AATGAAGAATAAACAACACTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1427
Db      ||||| 781 AATGAAGAATAAACAACACTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 840
Oy      ||||| 1428 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAATACGTT 1487
Db      ||||| 841 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAATACGTT 900
Oy      ||||| 1488 GATGTGATACCAAGAAATGCTTAAAGTGAAGACAGTCTTAAACAGTAAACGAACGTAAC 1547
Db      ||||| 901 GATGTGATACCAAGAAATGCTTAAAGTGAAGACAGTCTTAAACAGTAAACGAACGTAAC 960
Oy      ||||| 1548 TTAAGCTTCAAGATTTATACGATCCCTGATTAAGGCTTAACTACTCTCAACAATCTC 1607
Db      ||||| 961 TTAAGCTTCAAGATTTATACGATCCCTGATTAAGGCTTAACTACTCTCAACAATCTC 1020
Oy      ||||| 1608 GATGCTTTTGTATTAATGAACCTTAACTGAAGAAAGTAGAGATTAATCAAGATAC 1667
Db      ||||| 1021 GATGCTTTTGTATTAATGAACCTTAACTGAAGAAAGTAGAGATTAATCAAGATAC 1080
Oy      ||||| 1668 ACCAACCGTATCAATAACGTTTATATGGGCAAGCGACCCGAAGAGAAATGCTAGTAC 1727
Db      ||||| 1081 ACCAACCGTATCAATAACGTTTATATGGGCAAGCGACCCGAAGAGAAATGCTAGTAC 1140
Oy      ||||| 1728 CATTTAGCTGTGTGT 1742
Db      ||||| 1141 CATTTAGCTGTGTAT 1155

```

RESULT 7

```

US-10-631-558-6
; Sequence 6, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra

```

```

; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OR INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-631-558-6

```

```

Query Match      54.1% Score 1134.2; DB 7; Length 1327;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1142; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

Oy      ||||| 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAAACAGCAATGGTTGT 647
Db      ||||| 83 ATAGCTGCTCGTAATGGCTACTAGATGTCCTTCTGTAAATAACAGCAATGGTTGT 142
Oy      ||||| 648 AGCTTGTGTAATCTGTTGAGGGGACGAATCAAGACATTAATTTTGAATC 707
Db      ||||| 143 AGCTTGTGTAATCTGTTGAGGGGACGAATCAAGACATTAATTTTGAATC 202
Oy      ||||| 708 GATCTAATCAACGACCTGCTCATGAGGAAGAACAAGCAAGGCTTAAGTCAAAATCA 767
Db      ||||| 203 GATCTAATCAACGACCTGCTCATGAGGAAGAACAAGCAAGGCTTAAGTCAAAATCA 262
Oy      ||||| 768 AAACATTTGCTACTGTAATAGTGGCGCATGTCATATAAATTGAGAAAGCTGACTTA 827
Db      ||||| 263 AAACATTTGCTACTGTAATAGTGGCGCATGTCATATAAATTGAGAAAGCTGACTTA 322
Oy      ||||| 828 AAGCTATTCAGAAACATTTGATGCTTAACGTCACAGTAACGACTTGTAGGTC 887
Db      ||||| 323 AAGCTATTCAGAAACATTTGATGCTTAACGTCACAGTAACGACTTGTAGGTC 382
Oy      ||||| 888 ATTGATTTTCAAGCGATGCAACATTAATGATGGAAGGCAAGGCTACTTGTGTCAC 947
Db      ||||| 383 ATTGATTTTCAAGCGATGCAACATTAATGATGGAAGGCAAGGCTACTTGTGTCAC 442
Oy      ||||| 948 AAAGATGTTGCGTAACTTCCGACCAACCTGTCCAGAAATTTTGTAAAGGCAAT 1007
Db      ||||| 443 AAAGATGTTGCGTAACTTCCGACCAACCTGTCCAGAAATTTTGTAAAGGCAAT 502
Oy      ||||| 1008 GTGGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 1067
Db      ||||| 503 GTGGCGGTTTGAACCATATTAAGAAAAACAATCAAAAACCAAGGAAATCTGTTGATGTG 562
Oy      ||||| 1068 GAATATATCTGTACAGTTTCTCCCTTAAACCTGTGACGATTTTCAACAGGCTCAAA 1127
Db      ||||| 563 GAATATATCTGTACAGTTTCTCCCTTAAACCTGTGACGATTTTCAACAGGCTCAAA 622
Oy      ||||| 1128 GATATCTAAGCTATTGAAAAACATAGCTATGCGTGAACCATCACTCAAGAAATTACTA 1187
Db      ||||| 623 GATATCTAAGCTATTGAAAAACATAGCTATGCGTGAACCATCACTCAAGAAATTACTA 682
Oy      ||||| 1188 GCTCAAGCACAAAACATTTTAAACAAAACCAACCAGGCTATACGATTTTGAACGTCAC 1247
Db      ||||| 683 GCTCAAGCACAAAACATTTTAAACAAAACCAACCAGGCTATACGATTTTGAACGTCAC 742
Oy      ||||| 1248 TCCTCAATGCTCACTCATGACATGACATTTTCCGTAACGATTTTCAAAATGATCAAGAG 1307

```

|||||
Db 743 TCCTCATGTGTCATCATGACATGATGATTTCCGTACGATTTTACCAATGATCAAG 802
|||
Qy 1308 TTTACTACCGTGTAAATATGGGAAACAAGCTTATAGATTCATATATATATGCTG 1367
|||
Db 803 TTTACTACCGTGTAAATATGGGAAACAAGCTTATAGATTCATATATATATGCTG 862
|||
Qy 1368 AATGAAGAATTAACAACATGCTGATCTCTGAGAAATATATAGCTCTTAAATAGG 1427
|||
Db 863 AATGAAGAATTAACAACATGCTGATCTCTGAGAAATATATAGCTCTTAAATAGG 922
|||
Qy 1428 GAAAACCGTATGATCCCTTTGATGCGAGTCACTTGAATCTTCAATCAATAGCT 1487
|||
Db 923 GAAAACCGTATGATCCCTTTGATGCGAGTCACTTGAATCTTCAATCAATAGCT 982
|||
Qy 1488 GATGTCGATACCAACCAATTTGCTTAAATATGAGCACTCTTAAACAGTACGAACT 1547
|||
Db 983 GATGTCGATACCAACCAATTTGCTTAAATATGAGCACTCTTAAACAGTACGAACT 1042
|||
Qy 1548 TTAGACTTGAAGATTTATACGATCCTGCTGATAGGCTTAACTACTTACCAATCTC 1607
|||
Db 1043 TTAGACTTGAAGATTTATACGATCCTGCTGATAGGCTTAACTACTTACCAATCTC 1102
|||
Qy 1608 GATGCTTTGCTATATATGAGCACTATACCTTAACTGAAAATGAGATATCAAGATGAC 1667
|||
Db 1103 GATGCTTTGCTATATATGAGCACTATACCTTAACTGAAAATGAGATATCAAGATGAC 1162
|||
Qy 1668 ACCAACCGTATCATACCGTTATATGCGCAACGCAACCGGAGAGAGAAATGCTAGTAC 1727
|||
Db 1163 ACCAACCGTATCATACCGTTATATGCGCAACGCAACCGGAGAGAGAAATGCTAGTAC 1222
|||
Qy 1728 CATTTAGCTGNGGT 1742
|||
Db 1223 CATTTAGCTGNGGT 1237
|||

RESULT 8

US-10-631-558-3
; Sequence 3, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLON-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (777)
US-10-631-558-3

Query Match 15.7%; Score 330; DB 7; Length 777;

Best Local Similarity 100.0%; Pred. No. 4,3e-90;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1764 CCATAGCTGAGAGAGTGTGTTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 1823
|||
Db 448 CCATAGCTGAGAGAGTGTGTTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 507
|||
Qy 1824 TGGGAGAAAGCCCTACCAAGGCTGATGATGTGTGATTTGATCTTCCCTGGGAGAAAGCAGC 1883
|||
Db 508 TGGGAGAAAGCCCTACCAAGGCTGATGATGTGTGATTTGATCTTCCCTGGGAGAAAGCAGC 567
|||
Qy 1884 GAGCGATCATCTTGCACTTTAGAAATAGATGCAACGATGAGACACAAGGACATCTTAT 1943
|||
Db 568 GAGCGATCATCTTGCACTTTAGAAATAGATGCAACGATGAGACACAAGGACATCTTAT 627
|||
Qy 1944 AGAATTGAGACACCTGAGCAAGAGATATTCAGAAAACTGCTCCAGTGTATCTGC 2003
|||
Db 628 AGAATTGAGACACCTGAGCAAGAGATATTCAGAAAACTGCTCCAGTGTATCTGC 687
|||
Qy 2004 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTGTGTGAGACCAATCG 2063
|||
Db 688 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTGTGTGAGACCAATCG 747
|||
Qy 2064 AGCGATCTGCGCCCTTCAACCGATGTTCT 2093
|||
Db 748 AGCGATCTGCGCCCTTCAACCGATGTTCT 777
|||

RESULT 9

US-10-995-561-114
; Sequence 114, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-114

Query Match 15.7%; Score 330; DB 7; Length 2443;
Best Local Similarity 100.0%; Pred. No. 8,4e-90;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1764 CCATAGCTGAGAGAGTGTGTTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 1823
|||
Db 913 CCATAGCTGAGAGAGTGTGTTGATCATGCTGCTGGAGCTTCCATATGTGTGCGAGAAACG 972
|||
Qy 1824 TGGGAGAAAGCCCTACCAAGGCTGATGATGTGTGATTTGATCTTCCCTGGGAGAAAGCAGC 1883
|||
Db 973 TGGGAGAAAGCCCTACCAAGGCTGATGATGTGTGATTTGATCTTCCCTGGGAGAAAGCAGC 1032
|||
Qy 1884 GAGCGATCATCTTGCACTTTAGAAATAGATGCAACGATGAGACACAAGGACATCTTAT 1943
|||
Db 1033 GAGCGATCATCTTGCACTTTAGAAATAGATGCAACGATGAGACACAAGGACATCTTAT 1092
|||
Qy 1944 AGAATTGAGACACCTGAGCAAGAGATATTCAGAAAACTGCTCCAGTGTATCTGC 2003
|||
Db 1093 AGAATTGAGACACCTGAGCAAGAGATATTCAGAAAACTGCTCCAGTGTATCTGC 1152
|||
Qy 2004 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTGTGTGAGACCAATCG 2063
|||
Db 1153 ACAGGCAACGCGGAGAGAGTGAAGTGTGAGAGGACACCTGTGTGAGACCAATCG 1212
|||
Qy 2064 AGCGATCTGCGCCCTTCAACCGATGTTCT 2093
|||

```

Db      1213  AGCGATCTGGCCCTTACCGATGTTGCT 1242

RESULT 10
US-10-995-561-105
; Sequence 105, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-105

Query Match      15.7%; Score 330; DB 7; Length 2488;
Best Local Similarity 100.0%; Pred. No. 8.5e-90;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0,

QY      1764  CCCATAGCTGGAAGTGTGTTTGATCATCTGCTGGACCTTCTATGTGTGCGAAGAACG 1822
      |||||
DB      913  CCCATAGCTGGAAGTGTGTTTGATCATCTGCTGGACCTTCTATGTGTGCGAAGAACG 972

QY      1824  TGGGAGAGCCCTTACCAAGGCTGATGTAGTGAATTGACTTGGCTGGAGAGGACG 1883
      |||||
DB      973  TGGGAGAGCCCTTACCAAGGCTGATGTAGTGAATTGACTTGGCTGGAGAGGACG 1033

QY      1884  GAGCGCATCACTTGACCTTCTAGAAATGATGCAACGATCAGACACAGACATCTTAT 1943
      |||||
DB      1093  GAGCGCATCACTTGACCTTCTAGAAATGATGCAACGATCAGACACAGACATCTTAT 1092

QY      1944  AGAATTGAGACACCTGGAGCAGAAGATTAATCAGAGAAACCTGCTCAATGATCTGC 2003
      |||||
DB      1093  AGAATTGAGACACCTGGAGCAGAAGATTAATCAGAGAAACCTGCTCAATGATCTGC 1152

QY      2004  ACAGGCAACGGCCCGAGAGAGATGGAAGTGTGAGAGGCACTCTGTGTGACACATCG 2063
      |||||
DB      1153  ACAGGCAACGGCCCGAGAGAGTGGAAAGTGTGAGAGGCACTCTGTGTGACACATCG 1212

QY      2064  AGCGATCTGGCCCTTACCGATGTTGCT 2093
      |||||
DB      1213  AGCGATCTGGCCCTTACCGATGTTGCT 1242

RESULT 11
US-10-995-561-112
; Sequence 112, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 6510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-112

```

Query Match	15.7%	Score 330	DB 7	Length 6510
Best Local Similarity	100.0%	Pred. No. 1.5e-89		
Matches 330	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1764	CCCATAGCTGAGAAAGTGTTTGATCATGCTGTGGAACTTCTTATGATGTCGAGAAAC	18223	
Db	913	CCCATAGCTGAGAAAGTGTTTGATCATGCTGTGGAACTTCTTATGATGTCGAGAAAC	972	
QY	1824	TGGGGAAGCCCTTACCAAGGCTGATGATGTTGATTGCTTGGGAGAAAGCAGC	18833	
Db	973	TGGGGAAGCCCTTACCAAGGCTGATGATGTTGATTGCTTGGGAGAAAGCAGC	10323	
QY	1884	GSAGGCATCATCTGTGCATTTCTAGAATATGATGCAACGATTCAGAACACAAGACATCTTAT	19433	
Db	1033	GSAGGCATCATCTGTGCATTTCTAGAATATGATGCAACGATTCAGAACACAAGACATCTTAT	10923	
QY	1944	AGAAATTGGAGACACCTGTGAGCAGAGAAGATATCAGAGAAACCTGCTCAGTGCACTGC	20033	
Db	1093	AGAAATTGGAGACACCTGTGAGCAGAGAAGATATCAGAGAAACCTGCTCAGTGCACTGC	11533	
QY	2004	ACAGGCAACGCGCCGAGAGAGAGTGAAGTGTAGAGGCAACCTCTGTGCAGACCAATCG	20633	
Db	1153	ACAGGCAACGCGCCGAGAGAGAGTGAAGTGTAGAGGCAACCTCTGTGTGACAGCACATCG	12123	
QY	2064	AGCGGATCTGGCCCTTACACGATGTTGCT	2093	
Db	1213	AGCGGATCTGGCCCTTACACGATGTTGCT	1242	

```

RESULT 12
US-10-995-561-117
; Sequence 117, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117
; LENGTH: 7823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-117

Query Match      15.7%; Score 330; DB 7; Length 7823;
Beat Local Similarity 100.0%; Prid. No. 1.7e-89;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1764  CCCATAGCTGAGAGTGTGTTTGTATCANGCTCTGCGGACTTCCTATGTGTGTCGAGAAAGC 1823
        |||||
DB      913  CCCATAGCTGAGAGTGTGTTTGTATCANGCTCTGCGGACTTCCTATGTGTGTCGAGAAAGC 972

QY      1824  TGGAGGAAGCCCTTACCAAGGCTGGATGATGTGTAATTGTACTTCGCTGGAGAAAGCAAC 1883
        |||||
DB      973  TGGAGGAAGCCCTTACCAAGGCTGGATGATGTGTAATTGTACTTCGCTGGAGAAAGCAAC 1032

QY      1884  GAGAGCATCACTTGCACCTTCAGAAATATGATGCAAGATCAAGACACAAAGGACATCCAT 1943
        |||||
DB      1033  GAGAGCATCACTTGCACCTTCAGAAATATGATGCAAGATCAAGACACAAAGGACATCCAT 1092

QY      1944  AGAATTGGAGACACCTTGGAGCAGAGAGATATGAGAGAAACCTGTCTCAAGTCATCTGC 2003
        |||||
DB      1093  AGAATTGGAGACACCTTGGAGCAGAGAGATATGAGAGAAACCTGTCTCAAGTCATCTGC 1152

QY      2004  ACAGGCAACGGCCGAGAGAGATGGAATGTATAGAGGCAACACTCTTGTGTGCAACCAATCG 2063
        |||||
DB      1153  ACAGGCAACGGCCGAGAGAGATGGAATGTATAGAGGCAACACTCTTGTGTGCAACCAATCG 1212

```


Qy 2064 AGCGATCTGGCCCCCTTCACCGATGTTGCT 2093
|||
Db 1213 AGCGATCTGGCCCCCTTCACCGATGTTGCT 1242
|||

Search completed: February 1, 2006, 14:43:15
Job time : 387.523 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2006, 16:09:44 ; Search time 8612.84 Seconds
(without alignments)
11385.988 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 2096

Sequence: 1 cgaagaccatcatcgtgtt.....cctcacgcagtcgttag 2096

Scoring table: IDENTITY NTC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	330	15.7	451	1	AL603368 DKFZp686C
2	330	15.7	560	7	CN419585 170005318
3	330	15.7	621	7	CN482442 hg20d08.y
4	330	15.7	626	7	CN419611 170005313
5	330	15.7	737	1	AU140993 AU140993
6	330	15.7	765	3	BM715855 UI-B-EJ0-
7	330	15.7	861	1	AU140971 AU140971
8	330	15.7	7501	4	BC078656 Homo sapi
9	330	15.7	7501	4	BC100030 Homo sapi
10	330	15.7	7777	4	CR749316 Homo sapi
11	330	15.7	7868	4	HSW80653 Homo sapi
12	330	15.7	7885	4	CR749281 Homo sapi
13	330	15.7	8411	4	HSW80692 Homo sapi
14	329	15.7	465	1	AI095589 cb24a08.x
15	329	15.7	734	1	AU140802 AU140802
16	328.4	15.7	8121	4	CR749317 Homo sapi
17	326.8	15.6	548	6	CD613789 56022208J
18	326.8	15.6	551	6	CD613788 56022208J
19	325.2	15.5	560	2	AI743013 WGSa12.x
20	319	15.2	500	2	BP956982 OVA-NN114
21	318.2	15.2	547	2	BE009640 PM4-BN017
22	318	15.2	943	5	BX391752 BX391752

23	317	15.1	705	1	AU140789
24	315.4	15.0	739	1	AL706215
25	310	14.8	603	7	CN419658
26	307	14.6	414	1	AA492032
27	305	14.6	518	1	AI093548
28	304	14.5	7434	11	DQ039102
29	298	14.2	538	1	AL603599
30	297.4	14.2	496	2	BG945197
31	296.4	14.1	303	1	AA852090
32	287	13.7	427	7	CN419650
33	287	13.7	452	7	CN419649
34	281	13.4	605	2	CN419657
35	280	13.4	408	2	BG900107
36	277.8	13.3	580	3	BQ292415
37	274	13.1	495	2	BP956977
38	271.8	13.0	494	2	BP96365
39	268.6	12.8	484	2	B1058354
40	268	12.8	725	5	BQ574857
41	264	12.6	454	8	W46530
42	261.8	12.5	474	6	CB536952
43	256.8	12.3	645	8	DN123922
44	251.8	12.0	314	7	CN419631
45	249.4	11.9	899	5	BX327266

ALIGNMENTS

RESULT 1
AL603368 451 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp686C067.5, mRNA sequence.
DEFINITION DKFZp686C067.5, mRNA sequence.
ACCESSION AL603368
VERSION AL603368.1 GI:15166874
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Ansong, W., Witkner, U., Mewes, W., Weil, B. and Wiemann, S.
TITLE EST (Ansong, W., Witkner, U., Mewes, W., Weil, B. and Wiemann, S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the CDNA sequencing consortium of the
German Genome Project.
No BL sequence available.
This clone (DKFZp686C067) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcententrum, Hubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686C067"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="686 (synonym: hicc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 15.7%; Score 330; DB 1; Length 451;
Best Local Similarity 100.0%; Pred. No. 8.4e-84;

	Matches	330;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
OY	1764	CCCATAGCTGAGAAGTGTTTTATCATGTCTGGGACTTCCTATATGTGTGGAGAAACG	1822							
Db	32	CCCAATAGCTGAGAAATGTTTTGATCATCTCGTGGAATTCTTATGTGTGGAGAAACG	91							
OY	1824	TGGAGAAAGCCCTTACAAGGCTGATGTATGTGATTAATCTGTGCTGGAGAAAGGCAGC	1883							
Db	92	TGGAGAAAGCCCTTACAAGGCTGATGTATGTGATTAATCTGTGCTGGAGAAAGGCAGC	151							
OY	1884	GGAGCGATCACTTGTGACTTCTTAGAATAATGATGCAACGATTCAGACACACAAGACATCTTAT	1943							
Db	152	GGAGCGATCACTTGTGACTTCTTAGAATAATGATGCAACGATTCAGACACACAAGACATCTTAT	211							
OY	1944	AGAAATTGGAGACACCTTGAGCACAAAGAGATTAATCGAGGAAAACCTTGCTCAAGTCATCTGC	2003							
Db	212	AGAAATTGGAGACACCTTGAGCACAAAGAGATTAATCGAGGAAAACCTTGCTCAAGTCATCTGC	271							
OY	2004	ACAGGCAACCGGCCGAGAGACGTGGAAAGTGTGAGAGGACACACTCTGTGTCAAGCCACATCG	2063							
Db	272	ACAGGCAACCGGCCGAGAGAGGTGGAAGGTGTGAGAGGACACACTCTGTGTCAAGCCACATCG	331							
OY	2064	AGCGGATCTGGGCCCTTCAACCGCATTTTGT	2093							
Db	332	AGCGGATCTGGGCCCTTCAACCGCATTTTGT	361							

QY	1764	CCCATPACTGGAAGTGTTTTGATCAGCTGCTGGGACTTCCTATGCTGTGGGGAACG	1822
Db	209	CCCATPACTGGAAGTGTTTTGATCAGCTGCTGGGACTTCCTATGCTGTGGGGAACG	268
QY	1824	TGGAGAAGCCCTTACCAAGCTGGATGATGTGTAACTTGTCTTGTGGGGAAGGCAAGC	1883
Db	269	TGGAGAAGCCCTTACCAAGCTGGATGATGTGTAACTTGTCTTGTGGGGAAGGCAAGC	328
QY	1884	GGAAGCATCACTTGTCACTTCTTGAATATGATGCAACGATCAGACACAAGGACATTCCTAT	1943
Db	329	GGAAGCATCACTTGTCACTTCTTGAATATGATGCAACGATCAGACACAAGGACATTCCTAT	388
QY	1944	AGAAATTGGAGACACCTTGGAGCAAGAGGATATGAGGAAACTGCTTCACTGCATCTGC	2003
Db	389	AGAAATTGGAGACACCTTGGAGCAAGAGGATATGAGGAAACTGCTTCACTGCATCTGC	448
QY	2004	ACAGGCAACGGCCGACGAGAGATGGAAGTGTGAAGGCAACACTTGTGCAAGCAACACATCG	2063
Db	449	ACAGGCAACGGCCGACGAGAGATGGAAGTGTGAAGGCAACACTTGTGCAAGCAACACATCG	508
QY	2064	AGCGGATCTGGCCCTTCAACCGATGTTTGGT	2093
Db	509	AGCGGATCTGGCCCTTCAACCGATGTTTGGT	538

[illegible]

RESULT 3	CN482442	621 bp	mRNA	linear	EST 26-APR-2004
LOCUS	CN482442				
DEFINITION	hw20d08.v.1 Human primary human oopl pericytes. Unamplified (hw)				
ACCESSION	CN482442				
VERSION	CN482442.1				
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				

REFERENCE	1 (bases 1 to 560)
AUTHORS	Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murase, J., Fisk, G. J., Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)
PIUMED	15146197
COMMENT	Contact: Brandenberger R

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 621)	Teai, J.Y. and Wistow, G.	Expressed sequence tag analysis of cultured primary human ocular pericytes	Unpublished (2004)	Contact: Wistow G

```

230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 560 Std Error: 0.00.
Location/Qualifiers
1. .560

```

```

Plate: 20 row: d column: 08
Seq primer: M13R1 reverse primer (ABI)
location/Qualifiers
1. .621
FEATURES
source
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_1b="GRN EB"
/note="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p23), and H9 (p26) maintained in feeder-free
conditions."

```

ORIGIN

Query Match	15.7%;	Score 330;	DB 7;	Length 560;
Best Local Similarity	100.0%;	Pred. No. 9e-84;		
Matches 330;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTCATGTTCTAGATCGACGGGCGCGCC (T)15-3']. cDNA was

Query Match	15.7%	Score 330	DB 1	Length 737
Best Local Similarity	100.0%	Pred. No. 9.8e-84		
Matches 330	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1764	CCCATAGCTGAGAAAGTGTGTTGATCATCTGCTGGGACTTCTCTATGTTGTCGAGAAACG	18223	
Db	362	CCCATAGCTGAGAAAGTGTGTTGATCATCTGCTGGGACTTCTCTATGTTGTCGAGAAACG	421	
QY	1824	TGGGAGAAAGCCTTACCAAGGCTGATGATGTTGATCTTCTGCTGGGAGAAAGCAGC	18833	
Db	422	TGGGAGAAAGCCTTACCAAGGCTGATGATGTTGATCTTCTGCTGGGAGAAAGCAGC	481	
QY	1884	GGAGCATCATCTGTGACTTCTTGAATATGATGACCAAGTCAAGACACAAGACATCCAT	19433	
Db	482	GGAGCATCATCTGTGACTTCTTGAATATGATGACCAAGTCAAGACACAAGACATCCAT	541	
QY	1944	AGAAATGAGACACCTGAGAGCAAGAAGATATCGAGGAAACCTGCTCAGTGACATCGC	20030	
Db	542	AGAAATGAGACACCTGAGAGCAAGAAGATATCGAGGAAACCTGCTCAGTGACATCGC	601	
QY	2004	ACAGCAACCGCCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGACAGCCACATCG	20633	
Db	602	ACAGCAACCGCCGAGAGAGTGAAGTGTGAGAGGCACACCTCTGTGACAGCCACATCG	661	
QY	2064	AGCGGATCTGCGCCCTTCACCGAGTTCGT	2093	
Db	662	AGCGGATCTGCGCCCTTCACCGAGTTCGT	691	
RESULT 6				
BM715855				
LOCUS	BM715855	765 bp	mRNA	linear
DEFINITION	UI-E-Eu0-ahj-h-11-0-UI.r2 UI-E-Eu0 Homo sapiens cDNA clone			EST 28-FEB-2002
ACCESSION	UI-E-Eu0-ahj-h-11-0-UI 5', mRNA sequence.			
VERSION	BM715855			
KEYWORDS	BM715855.1 GI:19029113			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homnidae; Homo.			
REFERENCE	1 (bases 1 to 765)			
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene			
	discovery			
JOURNAL	Genome Res. 6 (9), 791-806 (1996)			
PUBMED	8889548			
COMMENT	Contact: Soares, MB			
	Coordinated Laboratory for Computational Genomics			
	University of Iowa			
	375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA			
	Tel: 319 335 8250			
	Fax: 319 335 9565			
	Email: bento-soares@iowa.edu			
	Tissue Procurement: Dr. Gregg Hageman			
	cDNA library preparation: Dr. M. Bento Soares, University of Iowa			
	cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa			
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa			
	Clone Distribution: Researchers may obtain clones from Research			
	Genetics (www.reagen.com).			
	Seq primer: M13 Reverse.			
FEATURES				
source	Location/Qualifiers			
	1..765			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="UI-E-Eu0-ahj-h-11-0-UI"			
	/tissue_type="fetal eyes, lens, eye anterior segment,			
	optic nerve, retina, Retina Foveal and Macular, RPS and			
	choroid"			
	/dev_stage="fetal and adult"			
	/lab_host="DH10B (life Technologies) (T1 phage resistant)"			

/clone lib="U1-B-E30"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 U1-B-E30 is a subcloned cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes,
 AGAATCAGA; lens, CGATTAGCA; eye anterior segment,
 AATGCCGAT; optic nerve, CCAATTAAGG; retina, CGCCG; Retina
 Foveal and Macular, GTCC; RPE and Choroid, ACCCA. This
 library was created for the program, Gene Discovery in the
 Visual System, supported by National Eye Institute (NEI)."

Query Match	15.7%	Score 330	DB 3	Length 765
Best Local Similarity	100.0%	Pred. No. 9.9e-84		
Matches 330	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1764	CCCATACCTGAGAGAGTGTGTTGATCATGCGCTGCTGGACCTTCTTATGTGTGTGAGAAACG	1823	
Db	182	CCCATACCTGAGAGAGTGTGTTGATCATGCGCTGCTGGACCTTCTTATGTGTGTGAGAAACG	241	
Qy	1824	TGGGAGAAAGCCCTTACCAAGGCTGGATGATGATGATTTGTACTTGCCTGGAGAAAGCAGC	1883	
Db	242	TGGGAGAAAGCCCTTACCAAGGCTGGATGATGATGATTTGTACTTGCCTGGAGAAAGCAGC	301	
Qy	1884	GAGCGATCATCTTGCACTTCTAGAAATGATGTCACGATCAGACACAGACATCTTAT	1943	
Db	302	GAGCGATCATCTTGCACTTCTAGAAATGATGTCACGATCAGACACAGACATCTTAT	361	
Qy	1944	AGAAATTGAGACACCTGTGAGACAGAAAGATGATGAGAAACCTGCTCCAGTGTGATCTGC	2003	
Db	362	AGAAATTGAGACACCTGTGAGACAGAAAGATGATGAGAAACCTGCTCCAGTGTGATCTGC	421	
Qy	2004	ACAAGGCAACGCGCGAGAGAGTGGAACTGTGAGAGGACACCTCTGTGACAGACACATCG	2063	
Db	422	ACAAGGCAACGCGCGAGAGAGTGGAACTGTGAGAGGACACCTCTGTGACAGACACATCG	481	
Qy	2064	AGCGGATCTTGCGCCCTTCAACCGATGTTTCGT	2093	
Db	482	AGCGGATCTTGCGCCCTTCAACCGATGTTTCGT	511	

/clone_lib="686 (synonym: hlc3). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiIb"
/dev_stage="adult"
/note="fibronectin 1, differentially spliced"
1..7777
/gene="DKFZp686K08164"
266..7069
/gene="DKFZp686K08164"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18171.1"
/db_xref="GI:51476362."
/translation="MLRGPGGLLAVLCLGTAIVSTGASKSKRAQOVQSPVA
VQSQKCYDNGKHQYQINQOMERTYLGNALVCTCYGSRFPNCSKREAEETFDKXT
GATYRVGDTYERKDSIMIDCTICIGARGISCTIARCHEGQSXYIGPTMRPHET
GGYMLECVLGNGKGEWTKPIAEKCPDHAAGSYVGEIWEKPYOGMMVDCICGE
GSGRTCTSRNRCNDODTRTSYRIGPTWSKDRGNILQCTICNGNGEKCRHITSY
QTTSSGSGPPTDRAAVYQPHQPPPTGHCYDGSVYVSVQMLTKGNQMLCT
CLANGVSCQETAVTQYTGNSNGPCVLPETVNGRTYSCITTEGRDGHLMCSTTSY
BODKYSFCTDHTVLVOTRGNSNGALCHPPLYNHNTYDCTSEGRDNMKCGTQ
NYDADQKFCPEMAAHEICTTEGVWYRGDQMDKOHDMNRCTGVNGKGEWTC
IAYSLRQDCIYDDITVYVNDTFHKBHEHMLNCTFCGGRGRCMDPVDCODSET
GTFYQIGDSMEKYHGRVQCYCGKIGEMHQPIQTYSSSGPPEVFTTEPSQDN
SHIQMNAQPSHISKYILKMRKNSVGRKKEATIGHLNYSYIKGLKPGVYEGQLI
S1QYGHQSVTRPDTFTTSTSTPVSNTVYGETTPSPVATSESYETIATASFVWS
VASADVSQFVREVELSEBDEPOYLDSTATSVNIPDLPRKXYIVNYQISEDE
OSLSTSQTTAPADPPDQVDDTSIVRSRQAPITGRIVYSPSEGSSTEL
NLPETANSYTLSDIQGVONYITIVAEENGESTPVIOQETTGTRSDVPSPRDQ
FREVYDKYTIKMTTPESATGYRVVIVPMLBEGQRIKRNTPFAETVATGSLPGYT
YFQVFAVSHGRESKPLTAQOTTKLAPNLQVNSTDVLRAWPRTVATGSLPGYT
VGLTRGQPROVNVGVSXKPLNLPQSESVYSVALKGNBSPKATVFTLQPG
S1PNTVETETTYITVTPAPRIGFKLGRSOGSBAKAGNBSGASIVSGLQPG
VYVYTIQVLRDQERDAPLVNKVYVPLSPNTLHLEANDTCVLTYSWERSPTPT
GYYITTPNGQAGNSLEVAHDOSSCTFDNLSPGLNVSYYTKYDKRESPIST
I1PEVQTLDSFVDTDSIGLRTPLNLSITIGRITVVAAGEG1PEDEVDSV
GYTYVGLBPGIDYD1SVTLINGESAPVLTQOVAPEPTDLRFTNIGPTMRVW
APPSIDLTNPLVRSVPVKNBDEVAELISPDNAVALNLTGTEVVSVSVEEH
ESTPLAGROKTDSPGTGIDFSDITANSFVHWIAPRATYGRIRHHPHSGRPE
DHPHSRNSITLNLTPGTEVYSIVANGRESPLIGQOSVSPVPRDLRYAATP
TSLISMDAPATVRYRTITGTEGANSPOQETTVGSKATITSGKREVDYITVY
AATGRDPSASSPISINRTEIDKPSQOVTVQONS1SVKMLPSSSPYGRVYTT
PNGRQPTKTKTAGPDDTENT1EGLQPTVEYVSVAGNSPESQPLVQAVTITPAP
TDLKFTQVPTSLAQMTPPNQLTGRVAVTPKEKTEPMKEINLAPDSSVVSGLM
VATKEVSUYALKDLTSRPAQGVTTLENSVSPRARVATDATETITL1SMRTETI
TGFOVDVAVANGOTPIORTIKDPVRSYITGLOPMDYKTYITLTDNARS5SVIDA
STAIAPSPURLATTPNLSLWSQPPARITTYIKTEKPSPREVPREPGVTE
AATTG1EPGTEYTIIVIALKNQKSPBLGRKKTQGBAL5QTTISWAPDQTBYSIIS
CHPVGDEBPLQFRVPGTSTATLGTGTGATYNI1VEALKDQQRHKBVEVTVGS
VNEG1NPDDSCFPDYTVSHAVGDEMREMSGFKLCOCLGFGSGHRECVSSRMC
HNGVNYKIGERKMDROGNGOMMSCTCLANGKGFPCDPHEATCYDDGTYHYGEMO
KEYLGALICCTCFGCGRGRCNCRPGRSGEPITPBGITGGSYNOYSORHYORTVNVNC
PIECFMPDVOADREDSRE"

ORIGIN

Query Match 15.7%; Score 330; DB 4; Length 7777;
Best Local Similarity 100.0%; Pred. No. 2.1e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1764 CCCATAGCTGAGAGTGTATTGATCATGCTGCTGGAGCTTCATGTGTCGAGAAACG 1823
Db CCCATAGCTGAGAGTGTATTGATCATGCTGCTGGAGCTTCATGTGTCGAGAAACG 865
1824 TGGGAGAGCCCTCAACAGGCTGATGATGATGATGATGATGATGATGATGATGAT 1883
Db TGGGAGAGCCCTCAACAGGCTGATGATGATGATGATGATGATGATGATGATGAT 925
1884 GGAACCATCTTGCATCTTGAATPAGATGCAAGATGAGACCAAGAGATCTTAT 1943
Db GGAACCATCTTGCATCTTGAATPAGATGCAAGATGAGACCAAGAGATCTTAT 985
1944 AGAATTGAGAGACCTGAGACCAAGAGATTAATCGAGAAACCTGCTCCAGTCATCTGC 2003
Db AGAATTGAGAGACCTGAGACCAAGAGATTAATCGAGAAACCTGCTCCAGTCATCTGC 1045

2004 ACAGGCAACGCGCCGAGAGAGTGAAGTGTGAGAGACACCTCTGTGACAGACCATCG 2063
1046 ACAGGCAACGCGCCGAGAGAGTGAAGTGTGAGAGACACCTCTGTGACAGACCATCG 1105
2064 AGCGATCTGAGCCCTTCACCGATGTTGCT 2093
1106 AGCGATCTGAGCCCTTCACCGATGTTGCT 1135

RESULT 11

HSMB06653
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL

COMMENT

This clone (DKFZp686M04163) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686M04163
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.
Location/Qualifiers

FEATURES

source

gene

CDS

/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="RZPD:DKFZp686M04163Q"
/db_xref="taxon:9606"
/clone="DKFZp686M04163"
/issue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlc3). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiIb"
/dev_stage="adult"
/note="fibronectin precursor"
1..7868
/gene="DKFZp686M04163"
267..7157
/gene="DKFZp686M04163"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA845714.1"
/db_xref="GI:34364617"
/db_xref="GCA:Q6N0A6"
/db_xref="InterPro:IPR000083"
/db_xref="InterPro:IPR000562"
/db_xref="InterPro:IPR002086"
/db_xref="InterPro:IPR003961"
/db_xref="InterPro:IPR006209"
/db_xref="UniProt/TREMBL:Q6N0A6"
/translation="MLRGPGGLLAVLCLGTAIVSTGASKSKRAQOVQSPVA
VQSQKCYDNGKHQYQINQOMERTYLGNALVCTCYGSRFPNCSKREAEETFDKXT
GATYRVGDTYERKDSIMIDCTICIGARGISCTIARCHEGQSXYIGPTMRPHET
GGYMLECVLGNGKGEWTKPIAEKCPDHAAGSYVGEIWEKPYOGMMVDCICGE
GSGRTCTSRNRCNDODTRTSYRIGPTWSKDRGNILQCTICNGNGEKCRHITSY
QTTSSGSGPPTDRAAVYQPHQPPPTGHCYDGSVYVSVQMLTKGNQMLCT

CLANGVSCOEIATVOTYGGNSNGEPCVLPTVNGRTFVSCCTEGRODHLMCSITSNY
BODKXSFCTDHTVLOTGNSNGALCHPPLVXNNHNTDSCBERLRNMKCGCTO
NYDADOKFCFMAAHEBICITNEGVNRIIGOWDKOHMGHMRCTVCGRNMGCTO
IAYSLRDOCIYDDITVYNNDRHKGHEGHLNCTCROGGRKWCDDVDOCDBET
CTFYIGDSMEKTYHGVYOCYCYGRTGSEMHQOPLQITPSSGPEVETITETPSPN
SHPIOMNAPQSHISKITILRMKPKSVGMKAKITPGLNHYTIKELKEVVEYEGLI
SIOQYHQBENRDEPTTSTPVSNTVETGTPSPPLVATSESTELTASFPVSM
VSADTVSGFVREYELSEBDEPOYLDLPSTATSVNI POLPKRXYVWYQISEGE
OSLITSTQTTAPDAPDPTVPODVTISVMSRPOARITGRIYVSPSVSGSTEL
NLPEASTVTSLOPQVOYNTIYAVENOSSTPVYQOETTGTRSTVSPRLO
FVEYDVKYITMTPTPESAVTGRVDIVNINPGEHGLPISRNTPAETVGLSEVY
YFKVFAVSHGRESKPLTAQOTTKLDAFNLDFVNETDSTVLRWMPRAQITGVL
VGLTRGQPROQNVGSPVSKYPLRNQPASEVTSVAIKGDSKAGVTTLOPG
SISPYNEVETITVITWTAPRIGFKLGVPSOGGABRETVSGSVVGLGR
VEYVITIOVLRDGOERDAPVKNVYVPLSPNMLHEAPDQVLTVMSEVTPIT
GRIITTPYNGQNSLEBVADSSCTFDNLSPLEYNVSYTYKDXKESVPISDT
IIPAVPEPTDLFTNIGPDMKVTWAPPSIDLNPLVSPKNEEDVAELISISD
NAVALJNLPGTEYVSVSVVEHESBTPLRGOKTGLDSPGIDPFSIDITANS
IAPRATITGYRIHHPHESGRPREDRVPHSNISITLNLFGTEYVSVIANGRE
SPILIGQOSTVSDVDRLEVNATPSTLISMDAPAVTVRYRITGETGNSVQEF
TVPSKSTATISGLKRGDYTYTVYAVTRGDSPASSKPSISINREIDKPSOMQYD
VQANSISYKMBESSPVYGYRTTTPKNGEPFKTKIAPDQTEMTIBLOQPTVEV
SVYAGNPSGESQPLVQTAVTTLIPAPDLKFTVYTPSLAOMTPPVVOLTRGVATP
KKTGPKRINLADPSSSVVSGLMWATKEVSVALKDITLSRPAQGVYTLNENP
PRARAVDATETITITSMRTKETITGFOVDAPVPAAGOTPIORTIKPVSYITGLQ
POTDKIYLTILNDNARSPPVIDASTAIDAPSNPLATPNSILVSMOPRARTIG
YIYKKGSPREYVPRPRPGVTRATITGLBPGTYTYVALAKNKGSEBPLIGKX
TDELPOLYTLPHNHLHGBEILDVSTVQKTPFVTHGYDNGIOLPGSGQPSVQ
OMIFEBHGFRTTPTTAPIRHPPRYPNVGEIIGHIPREDVDYLIYAGPLN
PNASTQEBLSQTTISWAFQDTSYIISCHVGEDEBLOQRPSTSTASLTGLTR
GATVNIIVEALKDORHKREEVTVGNSNEGILNOPDSCFDPYTHVYAVGDGE
RMSSEGFILCOCLRGSGHFCSSRWCHDNGVANKIGEKMDROENGOMMSCTG
NGKGEKDDPHATCYDDGKTYHVEGOKETYLGAICSTCTCGGGRGMRCDNCRPBG
EPTEGTTGOSTNOYSQRHQRKTNVNCPIECFMPDLVQADREDSR"

ORIGIN
Query Match 15.7%; Score 330; DB 4; Length 7868;
Best Local Similarity 100.0%; Pred. No. 2.1e-83;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1764 CCATAGCTGAGAAAGTGTTCATCATGCTGGAGACTTCTATGTCGAGAAACG 1823
Db 807 CCATAGCTGAGAAAGTGTTCATCATGCTGGAGACTTCTATGTCGAGAAACG 866
Qy 1824 TGGGGAAGCCTTACCAAGCGCTGATGATGATTAATTGATCTTCTGGAGAAAGCAGC 1883
Db 867 TGGGGAAGCCTTACCAAGCGCTGATGATGATTAATTGATCTTCTGGAGAAAGCAGC 926
Qy 1884 GAGCGCATCACTTGACCTTCTAGAAATGATGCAACGATCAGAGACAAGAGCATCTTAT 1943
Db 927 GAGCGCATCACTTGACCTTCTAGAAATGATGCAACGATCAGAGACAAGAGCATCTTAT 986
Qy 1944 AGAATTGAGACACTTGAGACAAGAGATAATTCAGAGAAACCTGCTCCAGTGCATCTGC 2003
Db 987 AGAATTGAGACACTTGAGACAAGAGATAATTCAGAGAAACCTGCTCCAGTGCATCTGC 1046
Qy 2004 ACAGGCAACGGCCGAGAGAGTGAAAGTGTGAGAGGACACCTCTGTGCAACCATATCG 2063
Db 1047 ACAGGCAACGGCCGAGAGAGTGAAAGTGTGAGAGGACACCTCTGTGCAACCATATCG 1106
Qy 2064 AGCGGATCTGGCCCTTACCGATGTTGCT 2093
Db 1107 AGCGGATCTGGCCCTTACCGATGTTGCT 1136

RESULT 12
LOCUS CR749281 7885 bp mRNA linear HTC 19-ANG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686f10164 (from clone DKFZp686f10164).
ACCESSION CR749281
VERSION CR749281.1 GI:51476291
KEYWORDS HTC.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE

1 (bases 1 to 7885)

AUTHORS

Koehler, K., Beyer, A., Mewes, H.W., Weill, B., Amid, C., Osanger, A.,
Rebo, G., Han, M. and Wiemann, S.

CONSTRM

The German cDNA Consortium

TITLE

Direct Submission

JOURNAL

Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764

COMMENT

Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp686f10164) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686f10164
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

Location/Qualifiers

1..7885

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp686f10164"

/issue_type="uterus endocheil, primary cell culture"

/clone_id="686 (synonym: hlc63). Vector pportil_sfi; host

DH10; sites SfiI + SfiIIp

/db_xref="GI:51476292"

/note="fibronectin 1, differentially spliced"

1..7885

/gene="DKFZp686f10164"

30..7103

/gene="DKFZp686f10164"

/cdon_start=1

/product="hypothetical protein"

/protein_id="CAH18136.1"

/db_xref="GI:51476292"

/translation="MAGPRRLCTCTGGGEGTPGAAGKRGPAATSLVLIAPVPPVP

FLTWPPSPMRQPPGGIRDRPRLRRRNIVATLCAPVASI.PHLNLMRGGGLL

LLAVLCLTRAVETGSKSKRQAOQVQVQPS.PAVASGKGCYDNGKQIYINOMERT

YLGNAVCTCYGSGRGNCSKPEALETCYDKTGNTRYGDTYRKHSDMIDCTV

KGEGRKISCTIARCHEGQSYKIGDTRWPHGTGTMLECYLGNGKMGKCPAL

KCFDHAAGTSYVGETHEKPYQGMVAVDCLCEGSGRLICTSRNCCNDQDTTSYAI

GDPTWSKKDNKGNLLOCICTGNGGEMKCEBRTSVQTSSSGSGFTDVAVYQOPHP

GGPPVHCYDTGSGVYVSGMOWLKTQGNOMLCTCICANGVSCOEIATVOTGNSNGE

PCVLPTVNGRTTSCATBGRDGHLMCTSTSYBODKTSFCTDHTVLOTGNSN

GAUCLHPPLVNNNHYDCTSEGRDNMKMGCTTQNDADQKFCFMAAHEBICITNE

GVMYRTGDOMDKOHDMHMRCTCVNGRGEWCIAYSQLRDOCIYDDITVYNNDRH

KHEBEGHMLNCTCFQGRGRMKCDPVDOCDSSTGFYVIGDSMEKHVGRVQCY

GGGIGEMHCOPLOTTPSSGPEVFTITRPSQNSHPIONNAPQSHISKILRMKPK

NSVGRKKEATPCHILNLSYTIKGLKGVVVEGOLISIOQYHQBENRDEPTTSTP

TSNTVGETTPSPVLVATSEVETITASFSVWVSASDVSGFRVETLSEBDEPO

YLDLPSTATSVNIPDLPKRXYVWYQISEDEBOSLILSTQTTAPDAPDPTVQY

DDTSIVRWSRPOAPITGYRIYVSPVSGSTELNIPETANSYTLISGLKGVVYNTI

YAVEENOSSTPVYIOETTGTPRSDTVSPRDLQFVETDKVTLIMTTPESAVTGR

VVYIPVNLPGHEHGLPISRNTPAETVGLNPLVSPKNEEDVAELISISPDNVLNML

LGTEYVSVSVVEHESBTPLRGOKTGLDSPGIDPFSIDITANSFPVWIMARATITGR

IRHHPHESGRPREDRVPHSNISITLNLFGTEYVSVIANGRESPILQOSVSDVDR

LEVNATPSTLISMDAPAVTVRYRITGETGNSVQEFTVPSKSTATISGLKGVVY

NTVYAGNPSGESQPLVQTAVTTLIPAPDLKFTVYTPSLAOMTPPVVOLTRGVATP

KKTGPKRINLADPSSSVVSGLMWATKEVSVALKDITLSRPAQGVYTLNENP

PRARAVDATETITITSMRTKETITGFOVDAPVPAAGOTPIORTIKPVSYITGLQ

POTDKIYLTILNDNARSPPVIDASTAIDAPSNPLATPNSILVSMOPRARTIG

YIYKKGSPREYVPRPRPGVTRATITGLBPGTYTYVALAKNKGSEBPLIGKX

TDELPOLYTLPHNHLHGBEILDVSTVQKTPFVTHGYDNGIOLPGSGQPSVQ

OMIFEBHGFRTTPTTAPIRHPPRYPNVGEIIGHIPREDVDYLIYAGPLN

PNASTQEBLSQTTISWAFQDTSYIISCHVGEDEBLOQRPSTSTASLTGLTR

QY 1764 CCCATAGCTGAGAAAGTGTTCATCATGCTGCTGGGACCTTCATATGTGTGGAGAAACG 1823
Db 808 CCCATAGCTGAGAAAGTGTTCATCATGCTGCTGGGACCTTCATATGTGTGGAGAAACG 867
QY 1824 TGGGAGAAAGCCCTACCAAGGCTGGATGATGTATGATTTGATCTTGGCTGGAGAAAGCAGC 1883
Db 868 TGGGAGAAAGCCCTACCAAGGCTGGATGATGTATGATTTGATCTTGGCTGGAGAAAGCAGC 927
QY 1884 GGACGCATCACTTGCACTTCTAGAAATAGATGACATCAGACACAAAGACATCTTAT 1943
Db 928 GGACGCATCACTTGCACTTCTAGAAATAGATGACATCAGACACAAAGACATCTTAT 987
QY 1944 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 2003
Db 988 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 1047
QY 2004 ACAGGCAACGCGCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 2063
Db 1048 ACAGGCAACGCGCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 1107
QY 2064 AGCGATCTGGCCCTTCAACCGATGTTGCT 2093
Db 1108 AGCGATCTGGCCCTTCAACCGATGTTGCT 1137

RESULT 14
AI095589/c 465 bp mRNA linear EST 05-OCT-1998
LOCUS
DEFINITION
QD2408.x1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone
IMAGE:1697174 3' similar to gb.X02761_cdel FIBRONECTIN PRECURSOR
(HUMAN); mRNA sequence.
ACCESSION
AI095589
VERSION
AI095589.1 GI:3434565
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 465)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1871 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amerham
High quality sequence stop: 446.
Location/Qualifiers
1. 465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1697174"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares pregnant uterus NbHPU"
/note="Organ: uterus, Vector: pT7T3-Pac; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo (dT) primer [5',
AACTGAGAATTCGCGCGCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."

ORIGIN
Query Match 15.7%; Score 329; DB 1; Length 465;

Best Local Similarity 99.7%; Pred. No. 1.6e-83;
Matches 329; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1764 CCCATAGCTGAGAAAGTGTTCATCATGCTGCTGGGACCTTCATATGTGTGGAGAAACG 1823
Db 434 CCCATAGCTGAGAAAGTGTTCATCATGCTGCTGGGACCTTCATATGTGTGGAGAAACG 375
QY 1824 TGGGAGAAAGCCCTACCAAGGCTGGATGATGTATGATTTGATCTTGGCTGGAGAAAGCAGC 1883
Db 374 TGGGAGAAAGCCCTACCAAGGCTGGATGATGTATGATTTGATCTTGGCTGGAGAAAGCAGC 315
QY 1884 GGACGCATCACTTGCACTTCTAGAAATAGATGACATCAGACACAAAGACATCTTAT 1943
Db 314 GGACGCATCACTTGCACTTCTAGAAATAGATGACATCAGACACAAAGACATCTTAT 255
QY 1944 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 2003
Db 254 AGAATTGGAGACACCTGAGACAAAGATATCGAGAAACCTGCTCCAGTCATCTGC 195
QY 2004 ACAGGCAACGCGCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 2063
Db 194 ACAGGCAACGCGCGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCAATCG 135
QY 2064 AGCGATCTGGCCCTTCAACCGATGTTGCT 2093
Db 134 AGCGATCTGGCCCTTCAACCGATGTTGCT 105

RESULT 15
AUI40802 734 bp mRNA linear EST 05-AUG-2002
LOCUS
DEFINITION
AUI40802 PLACE4 Homo sapiens cDNA clone PLACE4000266 5', mRNA
sequence.
ACCESSION
AUI40802
VERSION
AUI40802.1 GI:11002323
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 734)
Ota, T., Suzuki, Y., Itoh, K., Ishii, S., Yamamoto, J., Sugiyama, T.,
Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y. and Isogai, T.,
HRI human cDNA project (Ota, T., Suzuki, Y., Itoh, K., Ishii, S.,
Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
Masuho, Y., Isogai, T.)
Unpublished (2000)
JOURNAL
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1. 734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000266"
/clone_lib="PLACE4"
/clone_lib="PLACE4"
/note="Vector: pME18SFLU3"

ORIGIN
Query Match 15.7%; Score 329; DB 1; Length 734;
Best Local Similarity 99.7%; Pred. No. 1.9e-83;
Matches 329; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1764 CCCATAGCTGAGAGTGTGTTGATCATGCTGTGGGACTTCTATGTGTGAGAAACG 1823
    |||||||
Db 309 CCCATAGCTGAGAGTGTGTTGATCATGCTGTGGGACTTCTATGTGTGAGAAACG 368
    |||||||
QY 1824 TGGGAGAGCCCTTACCAAGCTGTGATGTGTGATTTGTACTTGCTGGGAGAAAGCAGC 1883
    |||||||
Db 369 TGGGAGAGCCCTTACCAAGCTGTGATGTGTGATTTGTACTTGCTGGGAGAAAGCAGC 428
    |||||||
QY 1884 GGAAGCATCATCTTGCACTTCTAGAAATGATGCAAGATGAGACAAAGAGCATCTTAT 1943
    |||||||
Db 429 GGAAGCATCATCTTGCACTTCTAGAAATGATGCAAGATGAGACAAAGAGCATCTTAT 488
    |||||||
QY 1944 AGAATTGAGAGACACTGTGAGCAAGAAAGATATCGAGAAAAGCTGCTCAAGTGCATCTGC 2003
    |||||||
Db 489 AGAATTGAGAGACACTGTGAGCAAGAAAGATATCGAGAAAAGCTGCTCAAGTGCATCTGC 548
    |||||||
QY 2004 ACAAGGCAAGGCGCGAGAGAGTGAAGTGTGAGAGGCAACTGTGTGACAGACCAATCG 2063
    |||||||
Db 549 ACAAGGCAAGGCGCGAGAGAGTGAAGTGTGAGAGGCAACTGTGTGACAGACCAATCG 608
    |||||||
QY 2064 AGCGGATCTGGCCCTTCAACCGATGTTGGT 2093
    |||||||
Db 609 AGCGGATCTGGCCCTTCAACCGATGTTGGT 638
    |||||||

```

Search completed: February 1, 2006, 12:22:51
 Job time : 8612.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:47:34 ; Search time 145.802 Seconds
(without alignments)
12632.698 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 3726

Sequence: 1 cgaagaccatcatgtgtgtt.....cctccacgatgtcgttag 2096

Scoring table: BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-x1p
-O=/cgn2.1/USFto_epool_p/US09940235/runcat_27012006_144217_27548/apd_query.fasta_1.7708
-DB=GeneSeq -QFMT=fastan -SUFFIX=tag -MINMATCH=0.1 -LOOPCD=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CGN_1_1_846 @runcat_27012006_144217_27548 -NCPU=6 -ICPU=3
-NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_GeneSeq_21.*
1: GeneSeq1980s.*
2: GeneSeq1990s.*
3: GeneSeq2000s.*
4: GeneSeq2001s.*
5: GeneSeq2002s.*
6: GeneSeq2003s.*
7: GeneSeq2003bs.*
8: GeneSeq2004s.*
9: GeneSeq2005s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1994	53.5	795	2	AA24797 Streptoki
2	1991	53.4	483	2	AA12885 Factor Xa
3	1984	53.2	413	8	AD192189 Streptoki
4	1984	53.2	414	2	AA10194 Streptoki
5	1984	53.2	414	2	AA63120 Streptoki
6	1984	53.2	414	2	AA24794 Streptoco
7	1984	53.2	414	2	AAW94664 Streptoco
8	1984	53.2	414	2	AA201556 Native st
9	1984	53.2	414	3	AA190282 S. equisi

10	1983	53.2	531	2	AA11829
11	1980	53.1	414	2	AAW94665
12	1978	53.1	483	2	AA12522
13	1976.5	53.0	1194	2	AAW21726
14	1969	52.8	499	2	AA12891
15	1966.5	52.8	813	2	AAW21728
16	1958	52.5	414	2	AAW86143
17	1958	52.5	414	2	AAW01295
18	1958	52.5	415	9	AE893425
19	1958	52.5	440	2	AA12889
20	1958	52.5	859	2	AA12893
21	1954	52.4	414	2	AA120202
22	1951.5	52.4	415	3	AA150870
23	1951.5	52.4	415	3	AA195953
24	1947.5	52.3	1194	2	AAW21725
25	1947	52.3	414	2	AAW86144
26	1947	52.3	414	3	AAW01296
27	1946.5	52.2	1194	2	AAW21724
28	1942	52.1	440	2	AAW06377
29	1928	51.7	413	2	AAW25020
30	1928	51.7	413	5	ABB80012
31	1928	51.7	413	6	ABG74199
32	1924	51.6	372	2	AA10200
33	1916	51.4	372	2	AA10197
34	1907.5	51.2	371	2	AA10195
35	1907	51.2	374	2	AA10198
36	1904	51.1	384	4	AAW74940
37	1902	51.0	384	3	AAW84007
38	1898	50.9	372	3	AAW84006
39	1897	50.9	800	2	AAW21723
40	1897	50.9	1181	2	AAW21727
41	1890.5	50.7	747	2	AA12894
42	1887	50.6	401	3	AAW84004
43	1887	50.6	413	3	AAW84005
44	1875	50.3	369	2	AA12892
45	1855	49.8	440	1	AAW50620

ALIGNMENTS

RESULT 1
AA24797
ID AA24797 standard; protein; 795 AA.

AC AA24797;
XX
DT 26-AUG-1999 (first entry)
XX
DE Streptokinase and maltose binding protein fusion protein.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nK; rSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
XX
PN MO9931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US026694.
XX
PR 15-DEC-1997; 97US-0069497P.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Reed GL;
XX
DR WPI; 1999-395183/33.
DR N-PSDB; AA80497.
XX

PD 27-JUN-1991.
 XX
 PF 07-DEC-1989; 89GB-00027722.
 XX
 PR 07-DEC-1989; 89GB-00027722.
 PR 07-DEC-1990; 90MO-GB001911.
 XX
 PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
 XX
 PI Dawson KM, Hunter MG, Czapliewsk LG;
 DR MPI: 1991-208151/28.
 DR N-PSDB; AAQ12162.
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 XX
 PS Disclosure; Page 96; 115pp; English.
 XX
 CC The protein is a recombinant product of a gene fusion construct. The
 CC sequence of the synthetic hirudin HV-1 genes was designed based on the
 CC published amino acid sequence (Dott J, et al FEBS Letters 165:180
 CC (1984)). The sequence of streptokinase was obtd. from PCR amplified
 CC chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
 CC used for the PCR were based on the pub- lished DNA sequence of S.
 CC equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
 CC 357-362 [1985]). The two sequences were used to construct an expression
 CC vector in which the hirudin gene is linked to the streptokinase gene via
 CC a linking sequence encoding a cleavage site for factor Xa.
 CC is present at the site of the target thrombus so the active agents are
 CC released specifically at the place where clot formation is occurring. See
 CC also AAR12887-R12889, AAR12891-R12894 and AAR12522. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX
 SQ Sequence 483 AA;

Alignment Scores:
 Pred. No.: 4,246-172 Length: 483
 Score: 1991.00 Matches: 403
 Percent Similarity: 84.29% Conservative: 10
 Best Local Similarity: 82.24% Mismatches: 27
 Query Match: 53.44% Indels: 50
 DB: 2 Gaps: 9

US-09-940-235-12 (1-2096) x AAR12885 (1-483)

QY 234 ACAGATTGACCATGCTGAGAGAGTGTTCATGTCGTGGACTTCATGTGCT 353
 |||||
 DB 4 ThrAspCysThrGluSerGlyGln-----AsnLeuCys-----LeuCys--- 16
 QY 334 CGAGAGAACTGGGAGAGGACGCGACGCATCATCTTCAAGAAATAGATGCAC 413
 |||||
 DB 17 -----GluGlySerAsn-----ValCysGlyGlnGlyAsnLysCys--- 28
 QY 414 GATCAGACACAAGACATCTATAGATTGGAGACCTCGAGCAAGAAGATATCGA 473
 |||||
 DB 29 -----IleLeuGlySerAspGlyGluLysAsn----- 37
 QY 474 GGAAGACCTGCTCAGTCATGTCACAGGCAACGCGAGAGAGTGAAGTGTGAGAG 533
 |||||
 DB 38 -----GlnCysVal---ThrGlyGluGly----- 44
 QY 534 CACACCTCTGTGCAACACATCGACCGGATCTGGCCCTTCACCGCATGTT----- 584
 |||||
 DB 45 -----ThrProLysProGlnSerHisAsnAspGlyAspPheGluLysLeuProGluGlu 62
 |||||
 QY 585 -----CGTATTCCTGGAAGCTGAGTGGCTGTCTTAAGCCGTCCATCTGTC 626
 |||||
 DB 63 TyrLeuGlnIleGluGlyArgIleAlaGlyProGluTyrLeuLeuAspArgProSerVal 82
 |||||
 QY 627 AACACAGAGCAATGGTGTGAGCGTGTGCTGCTGCTGTTGAGGGGACGAATCAAGCAT 686
 |||||

DB 83 AsnAsnSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIle 102
 QY 687 AGCTTAATATTTTGAATCGATCTAACATCAAGACCTGCTCATGAGAGAAAGCAGAG 746
 |||||
 DB 103 SerLeuLysPhePheGluIleAspLeuHisSerArgProAlaHisGlyGlySerGlu 122
 QY 747 CAAGGCTTAAGTCCAAAATCAAAAACATTTGCTACTGATAGTGCGCGCATGTCACTAAA 806
 |||||
 DB 123 GlnGlyLeuSerProLysSerLysPhePheIleThrAspSerGlyAlaMetProHisLys 142
 QY 807 CTTGAGAAAGCTGACTTAATCAAGCTTATTAAGACATTTGATGCTTAACCTGCACAGT 866
 |||||
 DB 143 LeuGlyLysValIleAspLeuLeuValAlaIleGlnGluGlnLeuIleAlaAsnValHisSer 162
 QY 867 AACGACGACTTCTTGAAGTCAATTTTGCAGCGCATGACCACTTACTGATCGCAAC 926
 |||||
 DB 163 AsnAspArgLysPheGluValIleAspPheIleSerAspAlaThrIleThrAspArgAsn 182
 QY 927 GGCAGAGTCTACTTTGCTGACAAAGATGTTGCGTAACTTGCCGACCACTGTCCAA 986
 |||||
 DB 183 GlyValValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGln 202
 QY 987 GAATTTTGTCTAAGCGACATGTGCGGCTTATGACCATATTAAGAAACCAATACAAAC 1046
 |||||
 DB 203 GluPheLeuLeuSerGlyHisValArgValArgProTyrLysGlyLysProIleGlnAsn 222
 QY 1047 CAAGCGAAATCTGTGATGAGAAATATATGTAATGTAAGTTTACTCCCTTAACCCGTAGAC 1106
 |||||
 DB 223 GlnAlaLysSerValAspValArgValGluTyrThrValGlnPheThrProLeuAspProAspArg 242
 QY 1107 GATTTTCAGACAGGCTCAAGAAATACATGATGATTTGAAAACATGAGCTATGCGTAGACAC 1166
 |||||
 DB 243 AspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThr 262
 QY 1167 ATCAGATCTCAAGAAATTAAGTCAAGCAACAAAGATTTTAAACAAACCAACCCAGGC 1226
 |||||
 DB 263 IleThrSerGlnGlnLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGly 282
 QY 1227 TATACGATTTATGAAGCTGACTCTCTCAATGCTCATGACATGACATTTTCCGTACG 1286
 |||||
 DB 283 TyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThr 302
 QY 1287 ATTTTACCAATGAGTCAAGAGTTTACTTACCGTGTAAATGGGAAACAAAGCTTATAGG 1346
 |||||
 DB 303 IleLeuProMetCysArgGlnGluPheThrThrHisValLysAsnArgGluGlnAlaTyrGlu 322
 QY 1347 ATCAATTAATAAATCTGCTGTAATGAAGAATAAACAACACTGACTGATCTGAGAAA 1406
 |||||
 DB 323 IleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLys 342
 QY 1407 TATTTACGCTCTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATGCGAGTCACTTAAA 1466
 |||||
 DB 343 TyrTyrValLeuLysLysGlyGlyLysProTyrAspProPheAspArgSerHisLeuLys 362
 QY 1467 CTGTTCACCAATCAAAATCGTTGATGTCATGATCAACCAATTTGCTAAAAAGTGGACACTC 1526
 |||||
 DB 363 LeuPheThrIleLysTyrValAspValAsnThrAsnGlnLeuLeuLysSerGluGlnLeu 382
 QY 1527 TTAACAGCTAGCAACGTAATCTAGACTTATGAGAGATTATACGATCTCGTATTAAGCT 1586
 |||||
 DB 383 LeuThrAlaSerGluAsnLeuAsnLeuAspPheArgAspLeuTyrAspProAlaAspLysAla 402
 QY 1587 AAACTACTTACCAACAATCTGATGCTTTTGGATTATGACATTAACCTTAATGAGAAA 1646
 |||||
 DB 403 LysLeuLeuTyrHisAsnLeuAspAlaPheGlyIleMetCysArgTyrThrLeuThrGlyLys 422
 QY 1647 GTAGAGATTAATACAGATGACCAACCGTATGATTAACGTTTATATAGGGCAAGCAACC 1706
 |||||
 DB 423 ValGluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgPro 442
 |||||
 QY 1707 GAAGAGAGAAATGCTAGCTACATTTAGCT 1736
 |||||
 DB 443 GluGlyGluAsnAlaSerTyrHisLeuAla 452

RESULT 3	
ID	ADL92189
ADL92189	standard; protein; 413 AA.
XX	
AC	ADL92189;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Streptokinase protein sequence.
XX	
KW	harvesting; recombinant; host cell; N-terminal leader peptide;
KW	pre-peptide; lantibiotic; post-translational modification;
KW	pharmaceuticals; vaccine; immunogenic.
XX	
OS	Unidentified.
XX	
PN	WO200309862-A1.
XX	
PD	04-DEC-2003.
XX	
PF	26-MAY-2003; 2003WO-NL000389.
XX	
PR	24-MAY-2002; 2002EP-00077060.
PR	07-FEB-2003; 2003US-00360101.
XX	
PA	(NANO-) APPLIED NANOSYSTEMS BV.
XX	
PI	Moll GN, Leenhouts CJ, Kuipers OP, Driessen AJM;
DR	WPI; 2004-042770/04.
XX	
PT	Harvesting a desired polypeptide produced by a recombinant host cell, for
PT	producing pharmaceuticals, comprises selecting a recombinant nucleic acid
PT	comprising nucleic acid fragments encoding a leader peptide and the
PT	polypeptide.
XX	
PS	Claim 4; Page 82-83; 109pp; English.
XX	
CC	The invention relates to a novel method for harvesting a (poly)peptide
CC	produced by a recombinant host cell. The novel method involves selecting
CC	a cell comprising a first nucleic acid encoding a leader peptide and a
CC	second nucleic acid fragment encoding the desired (poly)peptide. The
CC	first and second fragments are within the same open reading frame of the
CC	first nucleic acid and the leader peptide is functionally equivalent to
CC	an N-terminal leader peptide found with the pre-peptide of a lantibiotic.
CC	The host cells and nucleic acids are useful for producing, harvesting and
CC	post-translational modification of polypeptides. The polypeptides may be
CC	used in the production of pharmaceuticals, e.g. as antigen for vaccine or
CC	immunogenic composition. This sequence represents a polypeptide relating
CC	to the novel method of the invention.
XX	
SQ	Sequence 413 AA;
XX	
Alignment Scores:	
Pred. No.:	1.72e-171 Length: 413
Score:	1984.00 Matches: 383
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	53.25%
DB:	8 Mismatches: 0
	Indels: 0
	Gaps: 0
US-09-940-235-12 (1-2096) x ADL92189 (1-413)	
QY	588 ATTCGTCGACCTCGAGTGGCGCTGACCGCTCATCTGTCAACAGCCCAATTGGTTGTT 647
Db	1 lleaiaaglyproglintpneuueapargprosevalaasnaesergineuvalval 20
QY	648 AGCGTTCGCGTATCTGTGAGGGGAGCAATCAAGACATTAGCTTTAAATTTTGAATC 707
Db	21 SerValAlaIaGIYThrValaIglnGlyIrrnaSglnAspIleSerneuyspnepeglnlile 40
QY	708 GATCAACATCAAGCACTCTCAATGAGGAAGAAAGACAGAGCAAGCTTAAGTCCAAATCA 767

Db	41	AspLeuThrSerArgProIaHis6dIyGaiLystrGlnGlnGlyLeuSerProIySer	60
Qy	768	AAACCATTTGGTACGTATGATGAGCGCATGTGCATATAACTTGGAAAAAGCTGACTTACTA	827
Db	61	LysProPheAlaThrAspSerClyAlaMetSerHislybLeuGlnIlybAlaAspLeuLeu	80
Qy	828	AAGGCTATTCAAGAACAAATTGATGCTAACGCTCCACAGTAACGACGACTTCTTGAGGCTC	887
Db	81	LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspPtyrPheGlnVal	100
Qy	888	ATTGATTTTGGCAGAGTGCACCACTTTCGATCGAAACGGCAGAGTCTACTTTGGCTGAC	947
Db	101	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlybValTyPheAlaAsp	120
Qy	948	AAAGATGGTTGGTAACTTGCACCGCACCACCTCGCCAAAGAAATTTTGGCTAAGCGGACAT	1007
Db	121	LysAspGlySerValThrLeuProThrGlnProValGlnGlnIlybLeuLeuSerClyHis	140
Qy	1008	GTGGCGGTGAGCACATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG	1067
Db	141	ValArgValArgProIyGlyGlnbSerProIleGlnAsnGlnAlaIySerValAspVal	160
Qy	1068	GAATTAATCTGTACAGTTTACTTCCCTTAAACCTTGATGACATTTCAAGCCAGGTCTCAA	1127
Db	161	GluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuIlys	180
Qy	1128	GATACATAAGCTATTGAAAAACCTAGCTTCGGGTGACACATCAATCATCTCAAGAAATTA	1187
Db	181	AspThrIlybLeuLeuIystrLeuAlaIleGlyAspThrIleThrSerGlnIlybLeu	200
Qy	1188	GCTCAGACACAAGCATTTTAAACAAAAACCAACCGAGCTTATACGATTTATGAACGTGAC	1247
Db	201	AlaGlnAlaGlnSerIleLeuAlaAsnIyAsnHisProGlyTyThrIleTyGlnArgAsp	220
Qy	1248	TCCTCAATCGTCACTGATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG	1307
Db	221	SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuPromerAspGlnGln	240
Qy	1308	TTTACTTAACGGTGTAAAAAATCGGGAAACAACCTATATGAGTCAATAAAAAATCTGGCTG	1367
Db	241	PheThrTyraGValIyAsnArgGlnGlnAlaTyraGlyIleAsnIyIySerClyLeu	260
Qy	1368	AATGAAGAATAATAACAACACTGACCTGATCTCTGAGAAATTTTACGTCCTTAAAAAGGG	1427
Db	261	AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnbTyrrTyrrValLeuIyIyGly	280
Qy	1428	GAAGAACCGTATGATCCCTTGTATGCGACGTGACTTGAACCTGTTCCATCAATACGTT	1487
Db	281	GluIySPtyrAspProPheAspArgSerHisLeuIyIleuPheThrIleTystrVal	300
Qy	1488	GATGGCGCATCAACAGAAATGCTAAAAAGTAGAGAGCTCTTAACAGCTAGGAGAACGTAC	1547
Db	301	AspValAspThrAsnGlnLeuLeuIySerIyGlnIlybLeuThrAlaSerClybArgAsn	320
Qy	1548	TTAGACTTCAGAGATTTATACGATCTTCGTGATAGAGCTAAAGCTTACTTCAACAATCTC	1607
Db	321	LeuAspPheArgAspLeuTyraSPProArgAspIyAlaIyIleuLeuTyraAsnAsnLeu	340
Qy	1608	GATGCTTTGGTATTATGAGCTATACCTTAACTGAAAAGTAGAGAGATATCAGATGAC	1667
Db	341	AspAlaPheGlyIleMetAspTyrrThrLeuThrGlyIyIyValGlnAspAsnHisAspAsp	360
Qy	1668	ACCAACCGTATCAATACCGTTTATATGGGCAAGGACCCGGAAGAGAAAGCTAGCTAC	1727
Db	361	ThrAsnArgIleIleThrValIyTweGlyIyAsArgProGlnGlyGlnAsnAlaSerTyrr	380
Qy	1728	CATTAGCT 1736	
Db	381	HisIleuAla 383	

```

ID AAR10194 standard; protein; 414 AA.
XX
XX AAR10194;
XX
XX 28-MAR-1991 (first entry)
XX
XX Streplokinae encoded by synthetic gene.
XX
XX streptokinase; thrombolytic agent; myocardial infarction.
XX
XX Synthetic.
XX
XX EP407942-A.
XX
XX 16-JAN-1991.
XX
XX 11-JUL-1989; 89JP-00179432.
XX
XX 11-JUL-1989; 89JP-00179432.
XX
XX 27-NOV-1989; 89JP-00307957.
XX
XX 11-APR-1990; 90JP-00096830.
XX
XX (SAKA ) OTSUKA PHARM FACTOR.
XX
XX Maajima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
XX
XX WPI; 1991-016179/03.
XX
XX N-PSDB; AAQ10230.
XX
XX Synthetic gene encoding streptokinase - scale, high purity prodn. of
XX streptokinase used as a thrombolytic agent.
XX
XX Claim 1; Page 51; 76pp; English.
XX
XX Streptokinase and its derivatives can be produced in large quantities
XX with high purity for use as thrombolytic agents in patients with lung
XX thrombus or myocardial infarction. See also AAR10195-R10200
XX
XX Sequence 414 AA;
XX
Alignment Scores:
Pred. No.: 1.73e-171 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: Gaps: 0
US-09-940-235-12 (1-2096) x AAR10194 (1-414)
QY 588 ATTGCTGGAAGCTGAGTGGCTGCTAAGCCGTCATCTGTGAACAAGCCAAATTGGTTGTT 647
DB 1 TlelaaglyProgluTripleuleuAspArgProSerValAsnAspSerGlnLeuVal 20
QY 648 AGCGTTGCTGAGTCTGTGAGGGGAGCAATCAAGCATTAGCTTAAATTTTGGAAATC 707
DB 21 SerValaIaaglyThrValGlnGlyThraSnglnAspLleSerLeuysPheGlnLeu 40
QY 708 GATTAACATCAAGCACTGCTCATGAGAGAAAGACAGAGCAAGCTTAAGTCCAAATCA 767
DB 41 AspLeuThrSerArgProIaIhIseGlyGlySerThrGlnGlnGlyLeuSerProIysSer 60
QY 768 AAACCATTTGCTACTGATGATGGCGCGATGTCACATTAACCTTGAGAAAGCTGAATTACTA 827
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeu 80
QY 828 AAGGCTATTCAAGAAATTAATGCTTAACGTCCACAGTAAGAGAGACTTACTTGAGTTC 887
DB 81 LysAlaIleGlnGlnGlnLeuLleAlaAsnValHisSerAspAspTyrPheGlnVal 100
QY 888 ATTGATTTTGAACGATGAACCATTAAGTCAAGAAAGCGAAGGTCATCTTGCTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120

```

```

QY 948 AAGATGTTGCTAAGCTTGGCCGACCAACCTGTCCAAAGATTTTGTAAAGCGACAT 1007
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
QY 1008 GTGGCGCTTAGACCATTAAGAAAGAAACCAATCAAAACCAAGGAAATCTGTGATGTG 1067
DB 141 ValArgValArgProTyrLysGlnLysProIleGlnSnglnAlaLysSerValAspVal 160
QY 1068 GAATATATCTGATGAGTTTATCCCTTAAACCCGATAGACATTTTCAAGCCAGGCTCAA 1127
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1128 GATATCAAGCTAATTAAGAAACATAGTATCGGTGACACCATCATCATCTCAAGATTTACTA 1187
DB 181 AspThrLysLeuLeuLysSerThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 200
QY 1188 GCTCAAGCAAAAGCATTTTAAACAAAAACCAACCCAGGCTATAGCATTTATGAACGTGAC 1247
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 220
QY 1248 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1308 TTTACTTACCGGTGTTAAATAATCGGGAACAAAGCTTATAGGATCAATAAATAATCGTCTG 1367
DB 241 PheThrTyrArgValLysAsnAsnArgLysGlnAlaTyrArgLysLysSerGlyLeu 260
QY 1368 AATGAAGAAATTAACAACACTGACCTGATCTGTGAGAAATATTACGTCTTAAAAAGAG 1427
DB 261 AsnGlnGlnLleAsnAsnThrAspLeuLleSerGlnLysTyrTyrValLeuLysGly 280
QY 1428 GAAAAGCGGTATGATCCCTTGTATCGCATCTTGAACCTGTTCACCATTAATACGTT 1487
DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1488 GATGCGATACCAAGCAATGTCTAAABAAGGAGAGGCTTAAACAGCTAGAGAACGTGAC 1547
DB 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 320
QY 1548 TTAGACTTCAGAGATTTATACATCCCTGATATAGGCTAAAGCTTAACCTCAACAACTCTC 1607
DB 321 LeuAspPheArgAspLeuTyrAspProIaIhIseGlyGlyLysLeuLysLeuTyrAsnAsnLeu 340
QY 1608 GATGCTTTTGGTATTAATGACATTAACCTTAACCTGAAGAAAGTAGAGATTAATCAAGATGAC 1667
DB 341 AspAlaPheGlyLleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1668 ACCAACCCTATCATTAACCGTTTATATGGGAGAGACCCGAAAGAGAAATGCTAGCTAC 1727
DB 361 ThrAsnArgLleIleThrValTyrMetGlyLysArgProGlnGlnAlaSerTyr 380
QY 1728 CATTAAGCT 1736
DB 381 HisLeuAla 383

```

RESULT 5

ID AAR63120 standard; protein; 414 AA.

XX AAR63120;

XX 16-OCT-2003 (revised)

XX 25-MAR-2003 (revised)

XX 16-NOV-1994 (first entry)

XX Streptokinase.

XX Streptokinase; SK; Streptococcus equisimilis; plasminogen;

XX myocardial infarction.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers
FH Region 1..352
FT /note="claim 3, see CC"
FT Region 14..414
FT /note="claim 1, see CC"
FT Region 120..352
FT /note="claim 3, see CC"
FT Region 244..414
FT /note="claim 3, see CC"
FT Region 244..352
FT /note="claim 2, see CC"
XX MO9407992-A1.
XX 14-APR-1994.
XX 05-OCT-1993; 93WO-US009502.
XX 05-OCT-1992; 92US-00956692.
XX 29-SEP-1993; 93US-00128299.
XX (GEHO) GEN HOSPITAL CORP.
XX (HARD) HARVARD COLLEGE.
XX Reed GL;
XX WPI; 1994-135561/16.
XX DNA encoding a polypeptide which binds to plasminogen and corresponds to
PT region of streptokinase - useful to detect plasminogen in a sample and to
PT treat myocardial infarction.
XX Disclosure: Page 40-41; 62pp; English.
XX Nucleic acid comprising a sequence encoding amino acids 14-414 of
CC streptokinase, which binds to plasminogen and does not have a sequence
CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX Sequence 414 AA;
SQ
Alignment Scores:
Pred. No.: 1.73e-171 Length: 414
Score: 1984.00 Matches: 363
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-12 (1-2096) x AAR63120 (1-414)
QY 588 ATTGCTGACCTGATGCGTCTAGACCGTCCATCTGTCAACAACAGCCAAATTGTTGTT 647
Db 1 lleaaglyprrroglitrrpleuenuaspraproservalasnsbserglnleuval 20
QY 648 AGCGTGTGCTACTGTGAGGGGAGCAATCAAGCATTAGTCTTAAATTTTGAATC 707
Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuysphethegln 40
QY 708 GATCTAACATCCAGACCTGCTCATGAGAGAAAGACAGACAGCGCTTAAGTCCAAATCA 767
Db 41 AspleutrrSerarprrolahisgilyglyllystrnglulglnlyleuserProlysSer 60
QY 768 AAACCATTTGCTACGATGATGCGGAGATGTCACATTAACCTTGAGAAAGCTGACTACTA 827
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspleuLeu 80
QY 828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTAACAGCAGTACTTTGAGGTC 887
Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100

QY 888 ATTGATTTTTCAGACGATGCAACCATTAAGTATGAAACCGAAGGTCTACTTTGCTGAC 947
Db 101 lleaapPhehlaseraspalathrrillethrasparhsnsllyssValTyrPheAlaAsp 120
QY 948 AANGATGTTCCGTTAACCTTGCCGACCCAACTGTGTCAGAAATTTTGTCTAAGCGGACAT 1007
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSergLysHis 140
QY 1008 GTGCGGTTAGACCATTAATAAGAAAAACAATACAAACCAAGGAAATCGTTGATGTG 1067
Db 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATATACGTGTACAGTTTACTCCCTTAACCCCTGATGACGATTTACAGACGAGCTCAAA 1127
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 1128 GATACTAAGCTATTGAAAACATGACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSergInGlnLeuLeu 200
QY 1188 GCTCAAGCAACAAGCATTTTAAACAACCAACCCAGCGCTATACGATTTATGAACGTGAC 1247
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 220
QY 1248 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACAGATTTTACCAATGATCAAGAG 1307
Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240
QY 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGTGTG 1367
Db 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysSergLysLeu 260
QY 1368 AATGAAGAATAAACAACATGACCTGATCTGTGAGAAATATTCGCTCTTAATAAAAAAGG 1427
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSergLysTyrTyrValLeuLysGly 280
QY 1428 GAAAGCCGATATATCCCTTATGCGAGTCATGGAACCTGTTACCATCAATACGTT 1487
Db 281 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1488 GATGTCGATACCAACGAAATGTCTAAAAAGTACAGAGCTCTTAACAGCTACGAAACGTAC 1547
Db 301 AspValAspThrAsnGlnLeuLeuLysSergLysGlnLeuLeuThrAlaSerGlnArgAsn 320
QY 1548 TTAGACTTCAAGATTTATTCATCTCTGCTGATPAGGCTTAACACTCTTCAACAATCTC 1607
Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1608 GATGCTTTGTGATTTATGACCTATACCTTAACGTGAAAGATAGAGATTAATCAAGATAC 1667
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1668 ACCAACCGTATCATTAACCGTTTATATGCGCAACGACCCCAAGAGAGAAATGCTAGCTAC 1727
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnAlaSerTyr 380
QY 1728 CATTAGCT 1736
Db 381 HisLeuAla 383
RESULT 6
AA124794
ID AA124794 standard; protein; 414 AA.
XX
XX AA124794;
AC
XX
XX
DT 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX
XX
XX Streptococcus equisimilis native streptokinase.
DE
XX
XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW
KW nsk; bacterial; blood clot; thrombotic condition;

KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 PN MO9931247-A1.
 XX 24-JUN-1999.
 PD 15-DEC-1998; 98WO-US026694.
 PF 15-DEC-1997; 97US-0069497P.
 PR 15-DEC-1997; 97US-0069497P.
 XX (HARD) HARVARD COLLEGE.
 PA
 Read Gl;
 PI WPI, 1999-395183/33.
 DR N-PSDB; AAX80492.
 PT N-terminally deleted streptokinase.
 PS Claim 30; Page 60-61, 73pp; English.
 XX The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents native streptokinase (nsk). (Updated on 17-OCT-2003 to
 CC standardise OS field)
 CC
 CC
 XX Sequence 414. AA;
 SQ
 Alignment Scores:
 Pred. No.: 1.73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x AAY24794 (1-414)
 QY 588 ATTGCTGACCTGAGTGGCTGCTAGACGCTCATCTGTCAACAAGCAATTTGTTT 647
 Db 1 TlealaglyProglutirpeuleuaspargProseValashnserserIneuVal 20
 QY 648 AGCGTGTCTGACTGTGTGAGGGAAGATCAAGACATTAGTCTTAATTTTGAATC 707
 Db 21 SerValalaglyThrValGluGlyThrAsnGlnAspIleSerLeuysPheGluIle 40
 QY 708 GATCTAATCATGACGACTGCTCATGAGAAAGACAGACAGCTTAAGTCCAAATCA 767
 Db 41 AspLeuThrSerArgProAlahisGlyGlyThrGluGlnGlyLeuSerProlySer 60
 QY 768 AAACATTTGCTACGATAGTGGCGGAGTGTCACTAACTTGAAGAACTGACTTAC 827
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 LysPhehealathrasPserGlyValametserHislyseuGluGlyValAspLeu 80
 QY 828 AAGGCTATTCAGAAACATTCATGCTTACGCTCAAGTACGACGATCTTACGAGT 887
 Db 81 LysAlalileGlnGlnGlnLeuIleAlasnaHisserAsnAspArgPheGluVal 100
 QY 888 ATTGATTTTGCAGACGATGCAACCATCTAGATCGAAACGCGAAGTCTTACTTGTGAC 947
 Db 101 IleaspPhealaserAlathrIleThrAspArgAsnGlyValTyrPheAlasP 120
 QY 948 AAAGATGTTGCGTGAACCTTCCGACCAACCTGTCAGAAATTTTGTAAAGCGACAT 1007
 Db 121 LysaspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
 QY 1008 GTGCGGTTAGACCATTTAAAGAAACCAATACAAACCAAGGAAATCTGTGATGTG 1067
 Db 141 ValArgValArgProTyrGlyGluLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1068 GAATATACGTACAGTTACTCCCTTAAACCTGATGACGATTCAGACGAGTCTCAAA 1127
 Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1128 GATATACGTTATGAAAACATGACTGATGCTGATGACACCATGACATTCAGAAATTA 1187
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
 QY 1188 GCTCAACGACAAAGCATTTTAAACAAACCAACCAAGCTTACGATTTATGACGTCAC 1247
 Db 201 AlaglnIleGlnserIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 220
 QY 1248 TCCCTCAATCGTCACTCATGACAAATGACATTTTCCGTGAGATTTTACCAATGATCA 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
 QY 1308 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATGATCAATAAATAATCGTCTG 1367
 Db 241 PheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 1368 AATGAGAAATTAACAACATGACCTGATCTTGAGAAATTTACGTTCTTAAAAAGG 1427
 Db 261 AsnGluGlnIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLysGly 280
 QY 1428 GAAAGCGGATGATGCTTGAATGCGAGTCACTTGAACCTGTTCAACATCAATACGT 1487
 Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1488 GATGCGATACCAACGATTTGCTAAAGAGAGAGCTTTAAACGCTACGAAACGTAAC 1547
 Db 301 AspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuThrAlaserGluArgAsn 320
 QY 1548 TTAGACTTCAGAGATTTATTCGATCCTCGTATGAGGCTTAACTCTTCAACATCTC 1607
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1608 GATGCTTTGCTATTTATGACCTATACCTTAACTGAAAGATAGAGATTAATCAGATGAC 1667
 Db 341 AspAlaPheGlyIleLeuAspIyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
 QY 1668 ACCAAGCGATCAATACCGTTTATATGAGGAGAGACCGAGAGAGAAATGCTAGTAC 1727
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaserTyr 380
 QY 1728 CATTTAGCT 1736
 Db 381 HisLeuAla 383
 RESULT 7
 ID AAW94664 standard; protein; 414 AA.
 XX AAW94664;
 AC AAW94664;
 XX
 DT 17-OCT-2003 (revised)

DT 04-MAY-1999 (first entry)
 XX Streptococcus equisimilis native streptokinase.
 XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
 KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KW serine protease; fibrin; blood clot; thrombolytic;
 KW vascular thromboembolytic symptom; acute myocardial infarction;
 KW fibrinolysis; resistance.
 XX Streptococcus dysgalactiae subsp. equisimilis.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 XX
 XX US5876999-A.
 XX 02-MAR-1999.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 XX Wu H;
 XX WPI; 1999-189643/16.
 DR N-PSDB; AX16632.
 XX
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 PT chromolytic agent.
 XX
 XX Claim 1; Col 7-10; 17pp; English.
 XX
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence represents native SK. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HPIg) to
 CC plasmin (HPIa), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 1,73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x AA094664 (1-414)
 QY 588 ATTCGTGACCTGAGTGTGCTGTAGACCGCTCATCTGTCAACAACGCAATTTGGTTGTT 647
 Db 1 IleaIaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 648 AGCGTTGCTGATCTGTGAAGGGAAGCATCAAGATTAAGTTAAATTTTGAATC 707
 Db 21 SerValaIaGlyThrValaGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
 QY 708 GATCAATCATGACGACTGCTCATGAGAGAAAGACGAGAGAGGCTTAAGTCCAAATATCA 767
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnIleLeuSerProLysSer 60

QY 768 AACCATTTGCTACTGATAGTGGCGCATGTCATATAACTTGAAGAAAGTGAATCTACTA 827
 Db 61 LyeProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysValaIAspLeuLeu 80
 QY 828 AAGGCTATTCAAGAAACATTTGATGCTTAACGCTCAAGTAAAGACGACTTGTAGGTC 887
 Db 81 LysAlaIleGlnGlnGlnIleuLeuIleAlaAsnValHisSerAsnAspArgTyPheGluVal 100
 QY 888 ATTGATTTTGCACGCAATGCAACCATTAATCTGATGAAACGGCAAGGTCTACTTGTGAC 947
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyPheAlaAsp 120
 QY 948 AAGATGTTGCGGTAACCTTGGCGCAACCTGCTCAAGAAATTTTGTCTAAGCGGACAT 1007
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 1008 GTGCGCTGTAGACCATTAATAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTG 1067
 Db 141 ValArgValArgProTyLysGlnLysProIleGlnAsnGlnAlaLysSerValaIAspVal 160
 QY 1068 GAATATATCTGTACAGTTTACTCCCTTAACCTGATGACGATTTCAACGAGTCTCAA 1127
 Db 161 GluTyThrValaGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1128 GATACCTAGCTATTGAAGAACAATGACTGATGCTGATGACCAACATCATCTCAAGAATTACTA 1187
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 1188 GCTCAAGCACAAGACATTTTAAACAAACCAACCAAGCGCTATAGCAATTTATGAACGTGAC 1247
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrIleTyGlnArgAsp 220
 QY 1248 TCTCAATGCTCACTCAATGACATGACATTTTCCGATCGATTTTACCAATGATCAAG 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAlaSerGln 240
 QY 1308 TTTACTTACCGGTGTTAAATTCGGGAACAAGCTTATGAGATCAATTAATAATTCGTCGTG 1367
 Db 241 PheThrTyThrValaLysAsnArgGlnGlnAlaTyArgIleAsnLysLysSerGlyLeu 260
 QY 1368 AATGAAGAATTAACAACAACATGACCTGATCTGTGAGAAATTTACGTCCTTAAAAAGG 1427
 Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyTyValLeuLysLysGly 280
 QY 1428 GAAAGCCGATGATGCCCTTGTATGCGATGCTTGAACCTGTTACCATCAATATGCTT 1487
 Db 281 GluLysProTyThrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 300
 QY 1488 GATGCGATATCAACGAAATGCTTAAAAAGTGAAGCTTTTAAACGCTAGCGGAAGTAA 1547
 Db 301 AspValaAspThrAsnGlnIleLeuLysSerGlnGlnIleuLeuThrAlaSerGlnArgAsn 320
 QY 1548 TTAGACTTGAAGATTTATATGATCCGTCGATGAAGGTAAGCTTAACATCAATCTC 1607
 Db 321 LeuAspPheArgAspLeuTyAspProArgAspLysAlaLysLeuLeuTyAsnAsnLeu 340
 QY 1608 GATGCTTTGTATATATGACTATATCACTTATGGAAGAGTGAAGATTAATCAAGATAC 1667
 Db 341 AspAlaIleGlyIleMetAspTyThrLeuThrGlyLysValaGlnAspAsnHisAspAsp 360
 QY 1668 ACCAAGCGTATCATTAACCGTTTATATGAGGCAAGGCAAGCGGAAGAGAAATGCTAGTAC 1727
 Db 361 ThrAsnArgIleIleThrValTyMetGlyLysArgProGlnGlnGlnIleuLeuLysSerTyx 380
 QY 1728 CATTTAGCT 1736
 Db 381 HisLeuAla 383
 RESULT 8
 AA01556
 ID AA01556 standard; peptide: 414 AA.
 XX
 AC AA01556;

XX 17-OCT-2003 (revised)
 DT 18-JUN-1999 (first entry)
 XX
 DE Native streptokinase protein sequence.
 XX Antigenic peptide; streptokinase; streptokinase-specific antibody;
 KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
 KW rheumatic fever.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 PN WO908698-A1.
 XX
 PD 25-FEB-1999.
 XX
 PF 18-AUG-1998; 98WO-US017114.
 XX
 PR 18-AUG-1997; 97US-0055911P.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Reed GL, Parhami-Seren B;
 XX
 DR WPI; 1999-190113/16.
 XX
 PT New polypeptides which bind streptokinase-specific antibodies - useful in
 PT thrombolytic therapy.
 XX
 PS Disclosure; Page 12; 44p; English.
 XX
 CC The present sequence represents a native streptokinase. The specification
 CC describes a polypeptide which binds to a streptokinase-specific antibody
 CC and prevents the antibody binding to native streptokinase. The
 CC specification also describes a synthetic polypeptide (PI) comprising an
 CC epitope which binds to a streptokinase-specific antibody and reduces
 CC thrombolytic activity of streptokinase. PI is used in thrombolytic
 CC therapy, and to prevent or treat glomerulonephritis and rheumatic fever.
 CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 414 AA;

Alignment Scores:

Pred. No.: 1,73e-171 Length: 414
 Score: 1984.00 Matches: 383
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.25% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x AAY01556 (1-414)

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCATCTGTCAACAAGCCAAATTGGTTGT 647
 DB 1 IleaAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValVal 20
 QY 648 AGCGTTGCTGTAAGTGTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
 DB 21 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGlnIle 40
 QY 708 GATCTTAACATCAAGACCTGCTCATGGAGAAAGACAGACAGCTTAAGTCCAAATCA 767
 DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSer 60
 QY 768 AAACCATTTGCTACTGATGAGCGGAGTGCATAAATTGAGAAAGCTGAAGCTTA 827
 DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlyValAspLeuLeu 80
 QY 828 AAGGCTATTCAGAAACAATTGATCGCTTAAGCTCAGAGTAACAGCACTTCTTGAGGTC 887
 DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100

QY 888 ATTGATTTTGCAGAGCATGACCAACCAATTACTGATCGAAAGCGAAGTCACTTCTGAC 947
 DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
 QY 948 AAAGATGGTTCCGTGACCTTCCCGACCCAACTGTCTCAAGAAATTTTCTTAAGCGGACAT 1007
 DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 1008 GTGGCGGCTTAACCATATTAAGAAAGAAACCAATTAACAAGCAAGCAAAATCTGTGATG 1067
 DB 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1068 GAATATCTGTACAGTTTACTCCCTTAACCCCTGATGACGATTTGACAGCAAGTCTCAA 1127
 DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
 QY 1128 GATACCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTACTA 1187
 DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 1188 GCTCAGACCAACAAGCATTTTAAACAAAACCAACCGAGCTTATGATTTATGACGTGAC 1247
 DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysArg 220
 QY 1248 TCCTCAATCGTCATGACATGACATGTCATTTCCGTGATGATTTTACCAATGATCAAGAG 1307
 DB 221 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnIle 240
 QY 1308 TTTACTTAACCGTGTGTTAAAAATCGGAAACAACTTATAGATCAATTAAAAAATCGTCTG 1367
 DB 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
 QY 1368 AATGAAGAATTAACAACACTGACCTGATCTCTGAAATATTTACGTCCTTAAAAAGG 1427
 DB 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLeuGly 280
 QY 1428 GAAAGCGATGATGCTCTTGGATCGAGTCACTTGAACGTTGAACCATGAATTAACGTT 1487
 DB 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1488 GATGTCGATACCAACAAGATTGCTAAAAAGTAGACGACTTAAAGCTTAAGCGAAGCTAAC 1547
 DB 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnArgAsn 320
 QY 1548 TTAGACTTCAGAGATTATATAGATCTCTGTGATTAAGGCTTAAGCTTCAACAATCTC 1607
 DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLysTyrAsnAsnLeu 340
 QY 1608 GATGCTTTTGTATTAAGACTATACCTTAACTGGAAGAGTAGAGATTAATCAGATGAC 1667
 DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
 QY 1668 ACCAAGCGTATCATTAACGTTTATATAGGGCAAGGACCCGAAGAGGAATGCTAGCTAC 1727
 DB 361 ThrAsnArgIleIleIleThrValTyrMetGlyLysArgProGlnGlyLysAsnAlaSerTyr 380
 QY 1728 CATTAGCT 1736
 DB 381 HisLeuAla 383

RESULT 9
 AAY90282
 ID AAY90282 standard; protein; 414 AA.
 XX
 AC AAY90282;
 XX
 DT 12-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX
 DE S. equisimilis streptokinase.
 KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;

KM cardiovascular disorder; fibrinectin.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
FN BP1024192-A2.
XX
PD 02-AUG-2000.
XX
PE 23-DEC-1999; 99EP-00310541.
XX
PR 24-DEC-1998; 98IN-DE003825.
XX
PA (COUL) CSIR COUNCIL SCI IND RES.
XX
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
XX WPI; 2000-516032/47.
DR N-PSDB; AAA37633.
XX
PT Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
PS
XX
XX Example 3; Fig 3; 58pp; English.
XX
CC This sequence represents the human Streptococcus equisimilis
CC streptokinase protein sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 414 AA;

Alignment Scores:
Pred. No.: 1.73e-171 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-12 (1-2096) x AAY90282 (1-414)
QY 588 ATTGCTGACCTGAGTGGCTGCTAAGCCGCTCATCTGTCAACAAGCCAAATTGGTTGTT 647
DB 1 llaaiaaglyproglutripheuleuaspargprosevalaasnaserserleuvala 20
QY 648 AGCGTGTGGTAACTGTGAGGGGAGATCAAGATCAAGATTAGTTAAATTTTGGAAAC 707
DB 21 servaiaaaglythvalaiglucllythrasnglnasplleserleuyshepheglnuile 40
QY 708 GATCAACATCAACAAGCTGCTCATGAGAGAAAGAGAGCAAGAGCTTAAGTCCAAATCA 767
DB 41 Aspleuthrserargproalaahisgllygllysthrngluclnglyleuserprolysser 60

QY 768 AAACATTTGCTACTGATAGTGGCGAGATGTCAATTAACCTTGAGAAAGCTGACTTA 827
DB 61 lvsrprophelaathrspsersgllyalameSerHieylseuclulysalaaapleuleu 80
QY 828 AAGCTAATTCAGAACAAATTGATCGCTTAACGTCCACAGTAACGACTAATTGAGGTC 887
DB 81 lvsalaillelngluclnuileuilealaenvalahieserAenaspargtyrpegluval 100
QY 888 ATTGATTTTGCAGCGAGATGCAACCATTAATCGAATCGAAACGGCAAGGCTACTTTCGAC 947
DB 101 llaasphealaseraspalatrrillethraspargasnglylvsvaltyrpehalasp 120
QY 948 AAAGATGTTGGTAACTTGCAGCCAGCCAACTGTCGAAGAAATTTTGTAAAGGAGCAT 1007
DB 121 lvsaspeylservalthrleuprotthrnglnprovalglnluheuleusergllyhis 140
QY 1008 GTGGCGGTAGACCATATTAAGAAAAACAATACAAAACCAAGGAAATGTTGATGTG 1067
DB 141 ValargValargprotyrlysglylvsproilleglnasnglnalysSerValaspVal 160
QY 1068 GAATATACCTGTACAGTTTACTCCCTTAAACCTCGATGACGATTTTCAGACCGGCTCAAA 1127
DB 161 GluTyThrValGlnPheThrProleuAsnProAspAspAspPheArgProglYleuLys 180
QY 1128 GATACCTAAGCTATTGAAAACTAGCTATCGTGACACCATCAATCTCAAGAAATTACTA 1187
DB 181 AsprThrlyseuleuulsthrleuallaleglyAsprThrillethrsersglnuileu 200
QY 1188 GCTCAAGCAACAAGATTTTAAACAAACCAACCCAGGCTATACGATTTTGAACGTGAC 1247
DB 201 AlaGlnlaGlnserilleuAsnlyAsnHieProglYtyrThrilleTyrgluuArgAsp 220
QY 1248 TCTCATGTGCTACTGATGACATGACATTTTCCGACATTTTCCAAATGATCAAGAG 1307
DB 221 SerSerillevalThrHiaspAsnaspillePheArgThrilleuProMetAspGlnlu 240
QY 1308 TTACTCTACCGTGTAAAAATGGGAAACAAGCTTATAGATCAATAAAAATCGTGTG 1367
DB 241 PheThrTyArgVallylsasnasnargluGlnalatyArgilleAsnlyslsSerGlyLeu 260
QY 1368 AATGAAGAAATAAACAACACTGACTGATCTGAGAAATATTAACTGCTTAAAAAAGG 1427
DB 261 AsngluclnuileasnsnhrthraepulleleserGluyserytyrValleuylvsGly 280
QY 1428 GAAAGCCGTGATGCCCTTTGATGCGAGTCACTTGAACCTGTTACCAATCAATACGTT 1487
DB 281 GluTyProTyzaspProPheAspArgserHieuleuysleuPheThrilleystyVal 300
QY 1488 GATGTGATPACCAACAATTGCTTAAAGAGACAGCTTACAGCTAGCGAAGCGTAAC 1547
DB 301 AspValAsprThrAsngluLeuleuLysSerGluGlnleuThrAlaserGluArgAsn 320
QY 1548 TTGACTCGAGAGATTATATACATCTCGTATAGAGCTTAACTACTTCAACAATCTTC 1607
DB 321 leuAspPhearglspleuTyzAspProArgAspLyalaYslseuleuTyzAsnAsnleu 340
QY 1608 GATGCTTTTGGTATTAAGACTATACCTTAACGTGAAAAGATGAAGATTAACAGATGAC 1667
DB 341 AspAlaPheglYlleMetAsprtyrThrleuThrglylvsValGluAspAsnHiaspAsp 360
QY 1668 ACCAACCCTATCATACCGTTTATATGGGGAACGACCCGAAGAGAAATGCTAGGTAC 1727
DB 361 ThrAsnArgllellethThrValTyzMetGlyLysArgProgluGlnuAlaserTy 380
QY 1728 CATTTAGCT 1736
DB 381 HisleuAla 383

RESULT 10
ID AAR11829 standard; protein; 531 AA.
XX AAR11829;
AC AAR11829;

```

XX 08-JUL-1991 (first entry)
DT
XX FB-FB-SK fusion conjugate.
DE
XX Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KM streptokinase; fusion protein.
XX
XX Staphylococcus aureus.
OS
XX
FH Key Location/Qualifiers
FT Peptide 2..61
FT /label= FB monomer
FT Peptide 62..119
FT /label= FB monomer
FT Peptide 120..531
FT /label= streptokinase
XX
XX US5011686-A.
XX
XX 30-APR-1991.
XX
XX 15-NOV-1989; 89US-00437769.
XX
XX 21-SEP-1987; 87US-00099242.
XX
XX (CREA-) CREATIVE BIOMOLEC.
XX
XX Pang RHL;
XX
XX WPI, 1991-140198/19.
XX N-PSDB; AAQ11651.
XX
XX Impacting injectable fibrinolytic agent - with affinity for intravascular
PT thrombus, by linking agent to fibrin binding domain.
XX
XX Discloure; Fig 5; 18pp; English.
XX
XX The conjugate comprises an FB-FB dimer linked to streptokinase. The FB
CC fragment has selective affinity for fibrin, low affinity for fibrinogen,
CC and minimal immunogenicity, imparting thrombus- targeting capability.
CC See also AAR11821 and AAR11828
XX
XX Sequence 531 AA;
SQ
Alignment Scores:
Pred. No.: 2.37e-171 Length: 531
Score: 1983.00 Matches: 382
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.48% Mismatches: 0
Query Match: 53.22% Indels: 0
Gaps: 0
DB: 2
US-09-940-235-12 (1-2096) x AAR11829 (1-531)
QY 585 CGTATTCGTCGACCTGAGCTGCTGCTAACCCTGCTAATCTGTCAACAGCAACCAATGGTT 644
DB 117 LysIleAlaGlyProGluTrpLeuAspArgProSerValAsnAsnThrGlnLeuVal 136
QY 645 GTTACGCTTGGTACTAGTGTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGA 704
DB 137 ValSerValAlaGlyThrValAlaGluGlyThrAsnGlnAspIleSerLeuLysPhePheGlu 156
QY 705 ATCGATCTAATCAAGCACTGCTCATGAGAGAAAGAGAGCAAGCTTAAAGTCCAAA 764
DB 157 IleAspLeuThrSerArgProAlaHisGlyGlyLeuThrGlnGlnIleuSerProLys 176
QY 765 TCAAAACCAATTTGCTACTGATATGTCGCGGATGTCACATTAACCTTGAAGAAAGTGA 824
DB 177 SerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 196
QY 825 CTAAGGCTATTCAGAGACAAATTGATCGCTAAGTCCAGATGACAGCACTTCTTGGAG 884

```

```

DB 197 LeuLysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspArgTyPheGlu 216
QY 885 GTCATGATTTTTCAGACGATGCAACCATTAATGATGGAACGGCAAGGTACTTGGCT 944
DB 217 ValIleAspPheAlaSerAspAlaThrLileThrAspArgAsnGlyLysValTyPheAla 236
QY 945 GACAAAGATGGTTCGTACCTTGCAGACCCCAACCTGTCACAAATTTTCTTAAGCGGA 1004
DB 237 AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGly 256
QY 1005 CATGTCGCGTTAGACCATTAATAAGAAACCAATACAAACCAAGCAAAATCTGTGAT 1064
DB 257 HisValArgValArgProTyLysGluLysProIleGlnAsnGlnAlaLysSerValAsp 276
QY 1065 GTGGAATATATCTGATGAGTTTATCTCCCTTAAACCTGATGACGATTCAGACGAGTCTC 1124
DB 277 ValGluTyThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeu 296
QY 1125 AAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1184
DB 297 LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrLileThrSerGlnGluLeu 316
QY 1185 CTAGCTCAGACCAACGATTTTAAACAAAACCAAGGCTATACGATTTATGAACGT 1244
DB 317 LeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyThrLileTyGluArg 336
QY 1245 GACTCCCTCAATCGTCATCTGACATGACATTTTCCGTACGATTTTACCAATGATCA 1304
DB 337 AspSerSerLileValThrHisAspAsnAspIlePheArgThrLileuPromerAspGln 356
QY 1305 GAGTTTACTTACCGCTGTTAAAAATCGGAAACAAGCTTATAGATCAATAAAAATCTGGT 1364
DB 357 GluPheThrTyThrValLysValHisAsnArgGlnGlnAlaTyThrGluAsnLysSerGly 376
QY 1365 CTGAATGAAGAAATTAACAACACTGACCTGATCTCTGAGAAATTAAGTCTTAAAAA 1424
DB 377 LeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyTyValLeuLysLys 396
QY 1425 GGGGAAAAGCCGATATGATCCCTTTGATGCGATGCTTGAACCTTGCACCATGAATAC 1484
DB 397 GlyGluLysProTyThrAspProPheAspArgSerHisLeuLysLeuPheThrLileTyGly 416
QY 1485 GTTGATGTCATGCAACGAATGTCTAAAGTAGAGGCTTAAAGCTAGCGAAGCT 1544
DB 417 ValAspValAspThrAsnGluLeuLysSerGluIleuLeuThrAlaSerGluArg 436
QY 1545 AACTTACACTTCAAGATTTATACGATCTCTGATTAAGGCTAATCTTCAACAAT 1604
DB 437 AsnLeuAspPheArgAspLeuTyThrAspProArgAspLysAlaLysLeuLysTyAsnAsn 456
QY 1605 CTGATGCTTTTGGATTTATGACATATACCTTAACTGAAAAGTAGAGATTAATCAGAT 1664
DB 457 LeuAspAlaPheGlyLileMetAspTyThrLeuThrGlyLysValGluAspAsnHisAsp 476
QY 1665 GACACCAACCGTATCAATACCGTTTATATGGGAGAGGACCCGAAAGAGAAATGCTAAC 1724
DB 477 AspThrAsnArgLileThrValTyMetGlyLysArgProGluGluGluAsnLase 496
QY 1725 TACCATTTAGCT 1736
DB 497 TyThrIleuAla 500

```

RESULT 11

AAW94665

ID AAW94665 standard; protein; 414 AA.

XX AAW94665;

AC

XX 04-MAY-1999 (first entry)

DT

XX Streptococcus equisimilis mutant streptokinase K59B.

DE

XX Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;

KM plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
 KM serine protease; fibrin; blood clot; thrombolytic;
 KM vascular thrombolytic symptom; acute myocardial infarction;
 KM fibrinolysis; resistance.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Misc-difference 14 /note= "encoded by ACC"
 PN US5876999-A.
 XX
 PD 02-MAR-1999.
 XX
 XX
 XX 06-DEC-1995; 95US-00568393.
 XX
 PR 06-DEC-1995; 95US-00568393.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 PI Wu H;
 XX
 DR WPI: 1999-189643/16.
 DR N-PSDB; MAX16633.
 PT Mutant streptokinase polypeptide - useful as plasmin-resistant
 XX thrombolytic agent.
 PS Claim 4; Col 11-14; 17pp; English.
 XX
 CC The present invention describes a mutant streptokinase (SK) polypeptide
 CC in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
 CC segment of the corresponding native SK is replaced by another amino acid.
 CC The present sequence is mutant SK K59E. SK is a secretory protein of
 CC haemolytic Streptococcus able to activate human plasminogen (HpiG) to
 CC plasmin (HpiM), which is a serine protease able to catalyse the
 CC hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
 CC agent in the treatment of vascular thromboembolytic symptoms such as
 CC acute myocardial infarction. Compared with wild-type SK, the K59E mutant
 CC is more resistant to degradation by human plasmin and is more effective
 CC both in acting as a fibrinolytic agent and in activating human plasminogen
 CC
 XX
 SQ Sequence 414 AA;
 Alignment Scores:
 Pred. No.: 3,996-171 Length: 414
 Score: 1980.00 Matches: 382
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.74% Mismatches: 0
 Query Match: 53.14% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x AAM94665 (1-414)
 QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATGTCACACAGCCCAATTGGTTGTT 647
 Db 1 IleaIaGIyProGluTrpLeuLeuAspArgProSerValaAsnAsnSerGlnLeuValVal 20
 QY 648 AGCGTGTGCTGATCTGTTGAGGGGACGAATCAAGACATTAGCTTAAATTTTGAATC 707
 Db 21 SerValaIaGIyThValGIuGIyThraSngInaAspIleSerLeuAspHeGluIle 40
 QY 708 GATCTAACATCAAGCACTGCTCATGAGAGAAAGACAGAGGCTTAAGTCAAAATCA 767
 Db 41 AspLeuThrSerArgProIaHisGIyLysTrnGIuGIyLysLeuSerProGluSer 60
 QY 768 AAACCAATTGGTACTGATAGTGGCGGATGTCACTAACTTGAAGAACTGACTTACTA 827
 Db 61 LysProPhealIaThrAspSerGIyAlaMetSerHisLysLeuGIuLysAlaAspLeu 80
 QY 828 AAGCTATTCAAGAAACAATTGATCGCTAACGTCACAGTAAGACGACTTGTGAGGTC 887

Db 81 LysAlaIleGIuGIuGIuLeuIleAlaLeuValHisSerAsnAspArgTrpPheGIuVal 100
 QY 888 ATTGATTTTGGCAAGCAGATGACCACTTACTGATGAAACGGCAAGGCTTACTTGGTGCAC 947
 Db 101 IleAspPheAlaSerAspAlaThrIleTrnAspArgAsnGIyValItyrPheAlaAsp 120
 QY 948 AAAGATGGTTCGGTAACTTGGCCGACCCCAACCTGTCACAAATTTTGTAAAGCGGACAT 1007
 Db 121 LysAspGIySerValThrLeuProTrnGlnProValGIuGIuLeuLeuSerGIyHis 140
 QY 1008 GTGCGCGTTAGACCATATAAAGAAAAACCAATCAAAACCAAGCAAGAAATGCTGATGTG 1067
 Db 141 ValArgValArgProItyrLysGIuLysProIleGlnAsnGIuAlaLysSerValAspVal 160
 QY 1068 GAATATAGCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCAAGCTTCAAA 1127
 Db 161 GluTyTrThrValGIuGlnPheThrProLeuAsnProAspAspPheArgProGIyLeuLys 180
 QY 1128 GATACTAGCTATTGAAAACACTAGCTATTCGGTGACACCATCATCTCAAGAAATTACTA 1187
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGIyAspThrIleThrSerGIuGIuLeu 200
 QY 1188 GCTCAAGCAAAAGCAATTTTAAACAAAACCAACCAAGGCTATACGATTTTATGAAACGTGAC 1247
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisAspGIyTyTrThrIleTyrgIuArgAsp 220
 QY 1248 TCCTCAATCGTCACTCATGACCAATGCAATTTTCCGTACGATTTTACCAATGATCAAGAC 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGIuGIu 240
 QY 1308 TTTACTTACCGTGTAAATAATCGGAAACAAAGCTTATAGATCAATATAAATCTGCTGTG 1367
 Db 241 PheThrTyArgValLysAsnArgGIuGIuAlaTyArgIleAsnLysLysSerGIyLeu 260
 QY 1368 AATGAAGAATAAACAACACTGACCTGATCTGTGAGAAATATTACGTCCTTAAAAAAAGG 1427
 Db 261 AsnGIuGIuIleAsnAsnThrAspLeuIleSerGIuLysTyTrValLeuLysLysGIy 280
 QY 1428 GAAACCGATAGATCCCTTGTATGCGAGTCACTTGAACCTGTCCACATCAAAATACGT 1487
 Db 281 GluLysProTyArgProPheAspArgSerHisLeuLysLeuHeThrIleLysTyVal 300
 QY 1488 GATGTGATACCAACGAATTTGCTTAAATAAGTACAGAGCTTTAACAGCTACGGAACGTAC 1547
 Db 301 AspValAspThrAsnGIuLeuLeuLysSerGIuGIuLeuThrIleAspGIuArgAsn 320
 QY 1548 TTTAGATTCAGAGATTTATACGATTCCTCGATAGAGGCTTAACTACTCTCAACAATCTC 1607
 Db 321 LeuAspPheArgAspLeuTyArgProArgAspLysAlaLysLeuLeuTyTrAsnAsnLeu 340
 QY 1608 GATGCTTTTGTATTTATGACATTAACCTTAACTGAGAAAGTAAAGATTAATTCAGATAC 1667
 Db 341 AspAlaPheGIyIleMetAspTyTrThrLeuThrGIyLysValGIuAspAsnHisAspAsp 360
 QY 1728 CATTAGCT 1736
 Db 381 HisLeuAla 383
 RESULT 12
 AAR12522
 ID AAR12522 standard; protein; 483 AA.
 XX
 AC AAR12522;
 XX
 DT 25-MAR-2003 (revised)
 DT 17-SEP-1991 (first entry)
 XX
 DE Factor Xa-cleavable streptokinase-IBGR-hirudin.

XX Fusion protein; blood clotting; coagulation; fibrinolysis;
 KW antithrombotic; thrombolysis.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Protein 1..414
 FT Peptide /label= streptokinase
 FT /label= linker
 FT /note= "factor Xa cleavage site"
 FT Protein 419..483
 FT /label= hirudin
 PN W09109125-A.
 PD 27-JUN-1991.
 PF 07-DEC-1989; 89GB-00027722.
 XX 07-DEC-1989; 89GB-00027722.
 PR 07-DEC-1989; 89GB-00027722.
 PR 07-DEC-1990; 90MO-GB001911.
 PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
 PI Dawson KM, Hunter MG, Czaplewsk LG;
 XX WPI; 1991-208151/28.
 DR N-PSDB; AAQ12490.
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 XX
 XX Disclosure; Page 96; 115pp; English.
 PS
 CC The protein is a recombinant product of a gene fusion construct. The
 CC sequence of the synthetic hirudin HV-1 genes was designed based on the
 CC published amino acid sequence (Dodd J., et al FEBS Letters 165 180
 CC (1984)). The sequence of streptokinase was obtcd. from PCR amplified
 CC chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
 CC used for the PCR were based on the pub- lished DNA sequence of S.
 CC equisimilis strain HA6A (Malke, H., Roe, B., and Rerretti, J.J., Gene 34
 CC 357-362 [1985]). The two sequences were used to construct an expression
 CC vector in which the streptokinase gene is linked to the hirudin gene via
 CC a linking sequence encoding a cleavage site for factor Xa. The factor Xa
 CC is present at the site of the target thrombus so the active agents are
 CC released specifically at the place where clot formation is occurring. See
 CC also AAR12887-R12889; AAR12891-R12894 and AAR12885. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 CC
 XX
 SQ Sequence 483 AA;
 Alignment Scores:
 Pred. No.: 6,49e-171 Length: 483
 Score: 1978.00 Matches: 401
 Percent Similarity: 82.28% Conservative: 17
 Best Local Similarity: 78.94% Mismatches: 52
 Query Match: 53.09% Indels: 38
 DB: 2 Gaps: 8
 US-09-940-235-12 (1-2096) x AAR12522 (1-483)
 QY 588 ATTCGTGACCTGAGTGGCTGCTAGACCGTCATCTGTCAACAAAGCCAAATGTTGTT 647
 Db 1 Ileaaglyploglurleupneulphaproservalasnamserglnleuval 20
 QY 648 AGCGTGTGCTACTGTGTGAGGGAAGATCAAGACATTAGTCTTAATTTTGAATC 707
 Db 21 SerValaIaaglyThrValgluGlyThraanglnaapIleSerleuysPhepeglnlle 40
 QY 708 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGACAGAGCTTAAGTCCAAATCA 767

Db 41 AspleuThrSerargProalHisglYglYstrHnglGlnGlyLeuSerProlySer 60
 QY 768 AAACATTGCTACTGATAGTGGCGGCAATGTCACATTAACCTTGAGAAAGCTGACTTACTA 827
 Db 61 LysProPhealatrAspSergIyAlaMetProHisLysLeuGlnLysAlaAspleu 80
 QY 828 AAGCGTATTCAAGAACATTTGATGCTTAACGTCACAGTAAACGACGACTTGAAGTC 887
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspPlyrPheGlnVal 100
 QY 888 ATTGATTGTCAGACGATGCAACCATTAATGATCCGAAACGCGAAGTCTACTTGTGAC 947
 Db 101 IleAspPhealaserAspAlaThrIleThrAspArganglyLysValTyPheAlaAsp 120
 QY 948 AAAGATGTTGGTATCCTTCCGACCCACACTGTGTCAGAAATTTTGTAAAGCGACAT 1007
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
 QY 1008 GTGGCGTTGACCATATTAAGAAACCATATACAAACCAAGCAAAATCTGTGATGTG 1067
 Db 141 ValArgValArgProTyrglyglulysProlleGlnanglnAlaLysSerValAspVal 160
 QY 1068 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCAGTCTCAA 1127
 Db 161 GluTyThrValGlnPheThrProleuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1128 GATACCTAGCTATTGAAAACCTAGCTATCGGTGACACCATCACATCTCAAGAAATCTA 1187
 Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerglnLeuLeu 200
 QY 1188 GCTACAGACAAACATTTTAAACAAACCAACCGCTATAGATTATGAAAGCTGAC 1247
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrIsProGlyTyThrIleTyrglnArgAsp 220
 QY 1248 TCCCAATCGTCACCTGACACATGACATTTCCGTAGACATTTTACCAATGATCAAGAG 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnLys 240
 QY 1308 TTACTTACCGTGTGTTAAAAATCGGAAACAGACTTATAGATCAATTAATAAATCGTCTG 1367
 Db 241 PheThrTyrrHisValLysAsnArgGlnGlnAlaTyrglnIleAsnLysLysSergIyLeu 260
 QY 1368 AATGAAGAAATTAACAAACACTGACCTGATCTCTGAGAAATTTACGCTTTAAAAAGG 1427
 Db 261 AsnIleGlnIleLeuAsnThrAspLeuIleSergIlyTyTyTyValLeuLysLysGly 280
 QY 1428 GAAAGCGGTATGATCCCTTGAACGACGACCTGAAACGTTCAACATTAATGACCTT 1487
 Db 281 GluLysProTyArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTyVal 300
 QY 1488 GATGCGATACCAACGAATGCTTAAGAAAGTGAAGAGCTTTAAGCTAGGGAAGCTAAC 1547
 Db 301 AspValAsnThrArgnglnLeuLeuLysSergIyGlnLeuLeuThrAlaSerGlnArgAsn 320
 QY 1548 TTAGACTTCAGAGATTATACGATCTCGTATAGGCTAAACCTTACCAACATCTC 1607
 Db 321 LeuAspPheArgSerLeuTyArgProArgAspLysAlaLysLeuLeuTyTyAsnAsnLeu 340
 QY 1608 GATGCTTTGGTATTTAGACCTATACCTTAACTGAAAAGTGAAGATATACAGATGAC 1667
 Db 341 AspAlaPheGlyIleMetAspTyTyThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
 QY 1668 ACCAAGCGTATCATTAACGTTTATATGAGGCAAGGACCCGAAGAGAGAAATGCTAGCTAC 1727
 Db 361 ThrAsnArgIleIleThrValTyTMetGlyLysArgProGlnGlnLysAlaLysTyTy 380
 QY 1728 CATTACCTGTGTGTGCGACAGCGCAACGATTTGATCCCATAGCTGAAGATGTTTGAT 1787
 Db 381 HisLeuAla-----TyArgLysAspArgTyTyThrGlnGlnGlnValTyTySer 398
 QY 1788 CATGCTGCTGGAGCTTCTTATGTCGTCGAGAAAAGTGGAGAAAGCCCTTCAAGAGCTGG 1847
 ::: ::: ::: :::

Db 399 TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspIleGluGlyArg 418
 QY 1848 ATGATG---GTAGATTGACT-----TGCTGGAGAGAGCGAGCA 1886
 Db 419 ValValIYrThrAspCysThrGlnSerGlyGlnAsnLeuCysLeuGlySerAsn 438
 QY 1887 GCGATCACTTGACATCTTAGAAATAGATGCAACGATCAGACACAAAGACATCTTAGA 1946
 Db 439 -----ValCyGgIyGlnGlnIleAsnIleCys-----Ile 447
 QY 1947 ATTGGAGACACCTGGAGAGACAAAGATTAATCGAGAAACCTGCTCCAGTGCATTCGACA 2006
 Db 448 LeuGlySerAspGlyGlnIleCysAsn-----GlnCysVal---Thr 459
 QY 2007 GCGAAGCGCGCGAGAGAGTGAAGTGTGAGAGACACCTGTGGAGACCAATCGAGC 2066
 Db 460 GlyGluGly-----ThrpIolysProGlnSerHis 469
 QY 2067 GGATCTGGCCCTTCACCGATGTT 2090
 Db 470 AsnAspGlyAspPheGluGluIle 477
 RESULT 13
 ID AAM21726 standard; protein; 1194 AA.
 AC AAM21726;
 DT 17-OCT-2003 (revised)
 DT 01-OCT-1997 (first entry)
 XX Streptokinase/maltose binding protein fusion protein, RSK.
 DE Streptokinase/maltose binding protein fusion protein, RSK.
 KM Plasminogen-binding fragment; streptokinase; degradation; MBP;
 KM Chromolytic agent; blood clot; bolus; maltose-binding protein.
 OS Streptococcus dysgalactiae subsp. equisimilis.
 FH Key location/Qualifiers
 FT Protein 1..381
 FT /label= Maltose binding protein
 FT /note= "acts as blocking group"
 FT Protein 382..1194
 FT /label= Streptokinase
 XX MO9641883-A1.
 XX 27-DEC-1996.
 XX 07-JUN-1996; 96MO-US009640.
 XX 09-JUN-1995; 95US-00488940.
 XX (HARD) HARVARD COLLEGE.
 XX Reed GL;
 XX WPI; 1997-065469/06.
 XX Modified forms of streptokinase resistant to enzymatic cleavage - useful
 XX as thrombolytic agents in treating thrombosis and in medical equipment.
 XX Example 1; Page 11-12; 65pp; English.
 XX This sequence represents a fusion protein between maltose-binding protein
 XX and the plasminogen-binding fragment of streptokinase. This fusion
 XX protein was used in the design of modified streptokinase has an in vitro
 XX degradation rate at least 2 times slower than that of native
 XX streptokinase. Compounds containing modified streptokinases are
 XX specifically used as thrombolytic agents for dissolving blood clots in
 XX vivo in a mammal, preferably at a dose of 20000 U/kg, optionally as a
 XX bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
 XX standardise OS field)

XX SQ Sequence 1194 AA;
 Alignment Scores:
 Pred. No.: 1,31e-170 Length: 1194
 Score: 1976.50 Matches: 389
 Percent Similarity: 97.50% Conservative: 1
 Best Local Similarity: 97.25% Mismatches: 5
 Query Match: 53.05% Indels: 5
 DB: 2 Gaps: 3
 US-09-940-235-12 (1-2096) x AAM21726 (1-1194)
 QY 546 CAGACCAATCGAGC-----GGATCTGGCCCTTCACCGATGTTGTAATGCTGGA 596
 Db 366 GlnThrAsnSerSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
 QY 597 CTTGAGTGGCTGTAGACCGTGCATCTGTCAACAACGCAATGTTGTTAGCGTTGCT 656
 Db 385 ProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValIleSerValAla 404
 QY 657 GGTACTGTGAGGGAGACGAATCAAGACATTAGTCTTAATTTTGAATGATCTACA 716
 Db 405 GlyThrValGluGlyThrAsnGlnAspIleSerLeuIleAspPheGluIleAspLeuThr 424
 QY 717 TCAAGACCTGCTGATGAGAGAAAGACAGACGAGGCTTAAGTCCAAATCAAAACCAATT 776
 Db 425 SerArgProAlaHisGlyGlyLysTrpGlnGlnIleLeuSerProIleSerValPhePhe 444
 QY 777 GCTACTGATAGTGGCGGATGTGCATTAACCTTGAGAAAGCTGACTTAAGGCTAAT 836
 Db 445 AlaThrAspSerGlyAlaMetSerHisIleLeuGlnIleValAlaAspLeuLeuValAlaIle 464
 QY 837 CAAGACAATTGATTCGTTAAGTCCACAGTCAAGTCAAGTCACTTGAAGTCAATTTT 896
 Db 465 GlnGlnIleLeuIleAlaAsnValHisSerAsnAspAspIlePheGluValIleAspPhe 484
 QY 897 GCAAGCAATGCAACCAATTAATGATTCGAAGGCAAGCTCTTGGCTGCAAAAGATGCT 956
 Db 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValIlePheAlaAspLysAspGly 504
 QY 957 TCGGTAACTTGGCGGACCCAGCTGTCAGAAATTTTGGTAAAGCGGACATGTGCGGTT 1016
 Db 505 SerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHisValArgVal 524
 QY 1017 AGACATATAAGAAAAACCAATACAAACCAAGCAAGTCTGTGATGTGGAATTAAT 1076
 Db 525 Arg---TyrIleGlnIleProIleGlnAsnGlnAlaIleAspValAlaGluTyrThr 543
 QY 1077 GTACAGTTTACTCCCTTAAACCTGTAGACATTTTGACACGAGTCTCAAGATACATGAG 1136
 Db 544 ValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLysAspThrLys 563
 QY 1137 CTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATGATGATCAAGCA 1196
 Db 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleHisSerGlnGlnLeuLeuAlaGlnAla 583
 QY 1197 CAAGCATTTTAAACAAAAACCAACCAAGCTATACATTTATGAACGTGACTCTCAATC 1256
 Db 584 GlnSerIleLeuLeuLysAsnHisProGlyTyrThrIleTyrGluArgAspSerSerIle 603
 QY 1257 GTCACTCAAGACATGACATTTTCCGTACGATTTTACCATGATCAAGATTTACTTAC 1316
 Db 604 ValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyr 623
 QY 1317 CGTGTAAATAATCGGGAACCAAGCTTATAGATCAATTAATAATCGTGTGTAATGAAGA 1376
 Db 624 ArgValLysAsnArgGlnGlnAlaIleArgIleAsnLysSerGlyLeuAsnGlnGlu 643
 QY 1377 ATAAACAACATGACCTGATCTGTGAAATATATTCGCTTAAAAAAAGGGAAGCCG 1436
 Db 644 IleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGlyGluLysPro 663

QY 1437 TATGATCCCTTATGATGCGATCACTTGAAACTGTTCCATCAATAATGTTATGTCGAT 1496
Db 664 TTTAAAPProPhaSParSeriSleuValLeuPheThrIleuValAspValAsp 683
QY 1497 ACCAAGCAATGCTTAAAGAGCGCTCTTAAACAGCTGCGAAGCTTAAGCTTTC 1556
Db 684 ThrAspGluLeuLeuValSerGluGlnLeuThrAlaSerGluValAsnLeuAspPhe 703
QY 1557 AGAGATTATACGATCTCGTATTAAGGCTTAAGCTTAAGCTTAAGCTTTC 1616
Db 704 ArgAspLeuValThrAspProAlaAspValAlaValLeuLeuValThrAsnLeuAspPhe 723
QY 1617 GGTATTATGATTAACCTTAAGCTTAAGGAGAGATATATCAAGTGAACCAACCGT 1676
Db 724 GlyIleMetAspPyrThrLeuThrGlyValValGlnAspAsnHisAspPyrThrAsnArg 743
QY 1677 ATCATACGCTTATATGCGCAAGCAAGCCGAGAGAGAGAAATGCTAGCTTTCATTTGCT 1736
Db 744 IleIleThrValValThrMetGlyValArgProGluGlyGlnAsnAlaSerTyrHisLeuAla 763
RESULT 14
AAR12891
ID AAR12891 standard; protein; 499 AA.
XX AAR12891;
AC AAR12891;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-SEP-1991 (first entry)
XX Streptokinase fused to yeast alpha factor secretion sequence.
DE Streptokinase fused to yeast alpha factor secretion sequence.
XX Fusion protein; blood clotting; coagulation; fibrinolysis;
KM antithrombotic; thrombolysis; streptokinase.
XX Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 100009.
OS Streptococcus
XX Key
FH 1. 85 Location/Qualifiers
FT /label= pre-pro alpha factor secretion sequence
FT /note= "S. cerevisiae"
FT Protein 86..499
FT /label= mature streptokinase
XX MO9109125-A.
XX 27-JUN-1991.
XX 07-DEC-1989; 89GB-00027722.
XX 07-DEC-1989; 89GB-00027722.
PR 07-DEC-1990; 90MO-GB001911.
XX (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI Dawson KM, Hunter MG, Czapliewsk LG;
XX WPI; 1991-208151/28.
DR N-PSDB; AAQ12158.
XX Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT fractions having greater antithrombotic activity for therapy and
FT prophylaxis.
XX Disclosure; Page 86; 115bp; English.
XX The streptokinase sequence was obtd. from PCR amplified chromosomal DNA
CC from S. equisimilis (Lancefield's Gp C) ATCC 10009 or ATCC 9642 (the
CC primers used for the PCR were based on the published DNA sequence of S.
CC equisimilis strain H46 (Malke, H., Roe, B., and Perrecci, J.J., Gene 34
CC 357-362 [1985]). The gene was fused to DNA encoding the yeast alpha
CC factor pre-pro-secretion sequence in an expression vector, pSMD1/152, for
CC prodn. of recombinant strepto- kinase in S. cerevisiae strain BU2168. See

CC also AAR12887-R12889, AAR12892-R12894, AAR12885 and AAR12522. (Updated on
CC 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise
CC OS field)
XX
SQ Sequence 499 AA;
Alignment Scores:
Pred. No.: 4,36e-170 Length: 499
Score: 1969.00 Matches: 382
Percent Similarity: 96.00% Conservative: 2
Best Local Similarity: 95.50% Mismatches: 16
Query Match: 52.84% Indels: 0
Gaps: 0
US-09-940-235-12 (1-2096) x AAR12891 (1-499)
QY 537 ACCTCTGTGACAGACCATCGAGCGGATCTGGCCCTTACCGATGTTGTAATGCTGGA 596
Db 69 ThrIleAlaSerIleAlaAlaValGluGluGluValSerLeuAspValArgIleAlaGly 88
QY 597 CCTGAGTGGCTGTGATGACCCCTGATCTGTCACAAACAGCCAAATGTTGTTAGCGTCT 656
Db 89 ProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuValValSerValAla 108
QY 657 GGTACTGTGAGGGAAGCAATCAAGACTTATGCTTAAATTTTGAATGATCTAACA 716
Db 109 GlyThrValGluGlyThrAsnGlnAspIleSerLeuValSerPheGluIleAspLeuThr 128
QY 717 TCAGACCTGCTCATGAGGAGAAAGACAGAGAGGCTTAAGTCCAAATCAAAACATTT 776
Db 129 SerArgProAlaHisGlyGlyValThrGluGlnGlyLeuSerProValSerValProPhe 148
QY 777 GCTACTATAGTGGCGCGATGTCATTAATGAAAGCTGACTTAAGGCTATT 836
Db 149 AlaThrAspSerGlyAlaMetProHisLeuLeuValAspValAlaIle 168
QY 837 CAAGAACAAATGATCGGTACGTCACAGTAAGACGACTTGGATGATGATTTT 896
Db 169 GlnGluGlnLeuIleAlaAsnValHisSerAsnAspPyrPheGluValIleAspPhe 188
QY 897 GCAAGCGATGCAACCATTAAGTCAAGACGCAAGGCTCTTGGTGAACAAAGATGT 956
Db 189 AlaSerAspAlaThrIleThrAspArgAsnGlyValValIleAspValAspGly 208
QY 957 TCGGTAACTTGGCGAGCCAACTGTGCCAAGATTTTGTGAACGAGCATGTGCGGTT 1016
Db 209 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 228
QY 1017 AGACCATATTAAGAAAAACCAATACAAACCAACGAAATGTGTGATGTGGAATTAATCT 1076
Db 229 ArgProTyrLysGluValProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 248
QY 1077 GTACAGTTTACTCCCTTAAACCCGATGACGATTTGACAGCGGCTCAAAAGATACTAAG 1136
Db 249 ValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuValSerThrLys 268
QY 1137 CTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGATTAAGTCAAGCA 1196
Db 269 LeuLeuValThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuAlaGlnAla 288
QY 1197 CAAGACATTTTAAACAAAAACACCCAGGCTATTCGATTTATGAAGTACTCTCAATC 1256
Db 289 GlnSerIleLeuValSerThrHisAspProGlyTyrThrIleTyrGluValAspSerSerIle 308
QY 1257 GTACATCATGACATGATCATTTCCGTACGATTTTCAATGATGATCAAGGTTTACTTAC 1316
Db 309 ValThrHisAspAsnAspIlePheArgThrIleLeuProPheArgGlnGluPheThrTyr 328
QY 1317 CGTGTAAATAATCGGGAACAGCTTATAGATCAATATAAATCGTGTGATGAAGAA 1376
Db 329 HisValLysAsnArgGluGlnAlaTyrGluIleAsnValSerGlyLeuAsnGluGln 348
QY 1377 ATAAACAACACTGACTGATCTCTGAGAAAATATTAATGCTTAAAAAGGAGAAACCGC 1436

```

Db      |||
349  ILeuAaenThnAspLeuIleSerGIuLySTyTYrValleuLySGLyGIuLySPro 368
QY      1437 TATGATCCCTTGATGCGAGTCACTTGAAACCTTTCACATCAAAATACGTGATTCGAT 1496
Db      |||
369  TyAspProPheAspArgSerHisLeuLySLeuPheThrIleLySTyValAspValAa 388
QY      1497 ACCAAGCATTCGCTAAAGAGAGAGAGCTTAAACAGTAGCAAGCTAACTTAAGACTTC 1556
Db      |||
389  ThrAngIuLeuIleuLySerGIuInleuThrAlaSerGIuArgAsnLeuAspPhe 408
QY      1557 AGAGATTTATACGATCCTCGTATGAAGCTAAAGCTAAGTCAACAATCTCGATGCTTTT 1616
Db      |||
409  ArgAspLeuTyAspProArgAspLySAlaLySLeuLeuTyAsnAsnLeuAspAlaPhe 428
QY      1617 GGTATATGAGCTATACCTTAACCTGAAAGAGAGATATATCAAGATGACCAACCGT 1676
Db      |||
429  GlyIleMetAspTyThrLeuThrGIyLySValGIuAspAsnHisAspAspThrAsnArg 448
QY      1677 ATCATTAACGGTTTATATGAGGCAAGGAGACCGCAAGAGAGATGCTAGCTAATTAAGCT 1736
Db      |||
449  IleIleThrValTyMetGIyLySArgProGIuGIyGIuAsnAlaSerTyThrIleuAla 468

RESULT 15
AAW21728
XX      AAW21728 standard; protein; 813 AA.
AC      AAW21728;
XX      17-OCT-2003 (revised)
DT      01-OCT-1997 (first entry)
XX      Wild type plasminogen-binding fragment of Streptokinase.
DE      Plasminogen-binding fragment; streptokinase; degradation; MBP;
XX      thrombolytic agent; blood clot; bolus; maltose-binding protein.
XX      Streptococcus dysgalactiae subsp. equisimilis.
OS      MO9641883-A1.
XX      27-DEC-1996.
XX      07-JUN-1996; 96WO-US009640.
XX      09-JUN-1995; 95US-0048940.
XX      (HARD ) HARVARD COLLEGE.
XX      Reed GU;
XX      WPI; 1997-065469/06.
XX      Modified forms of streptokinase resistant to enzymatic cleavage - useful
PT      as thrombolytic agents in treating thrombosis and in medical equipment.
XX      Example 1; Page 12-13; 65pp; English.
XX      This sequence represents the wild type plasminogen-binding fragment of
XX      streptokinase. This fragment was used in the design of a modified
XX      streptokinase has an in vitro degradation rate at least 2 times slower
XX      than that of native streptokinase. Compounds containing modified
XX      streptokinases are specifically used as thrombolytic agents for
XX      dissolving blood clots in vivo in a mammal, preferably at a dose of 20000
XX      U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-
XX      OCT-2003 to standardise OS field)
XX      Sequence 813 AA;
XX      Alignment Scores:
XX      Pred. No.: 9,08e-170 Length: 813
XX      Score: 1966.50 Matches: 382
XX      Percent Similarity: 99.74% Conservative: 0

```

```

Best Local Similarity: 99.74% Mismatches: 0
Query Match: 52.78% Indels: 1
DB: 2 Gaps: 1
US-09-940-235-12 (1-2096) x AAW21728 (1-813)
QY      588 ATTGCTGACCTGAGTGGCTGTAGACCGTTCATCTGTCAACAACAGCAATGGTGTGTT 647
Db      |||
1  IleAlaGIyProGIuThrPheLeuAspArgProSerValAlaAsnSerGIuLeuVal 20
QY      648 AGCGTTGCTGTGATCTGTGAGGGAGCAATCAAGACATTAGCTTAAATTTTGAATC 707
Db      |||
21  SerValAlaGIyThrValGIuGIyThrAsnGIaAspIleSerLeuLySPhenGIuIle 40
QY      708 GATCTAACATCAGACTGCTGTATGAGAGAAACAAGAGAGAGAGAGAGAGAGAGAGAGAG 767
Db      |||
41  AspLeuThrSerArgProAlaHisGIyGIyLySThrGIuInGIyLeuSerProLySer 60
QY      768 AAACATTTGCTACTAGTAGTGGCGGATGTCATTAACCTTGAGAAAGCTTACTTA 827
Db      |||
61  LysProPheAlaThrAspSerGIyAlaMetSerHisLySLeuGIuLySAlaAspLeu 80
QY      828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCCACAGTAAACAAGCTTATGAGGTC 887
Db      |||
81  LysAlaIleGIuInGIuInleuIleAlaAsnValHisSerAsnAspAspTyThrPheGIuVal 100
QY      888 ATTGATTTTGCAGAGCGATGACACCATTTATCTGATCGAAACGGCAAGGTCTACTTGTGAC 947
Db      |||
101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGIySValTyThrPheAlaAsp 120
QY      948 AAGATGTTTCGGTAACTTGGCCAGCCCAACCTGTCCAAGATTTTGTGAAGCGGACAT 1007
Db      |||
121  LysAspGIySerValThrLeuProThrGlnProValGlnIuPheLeuSerGIyHis 140
QY      1008 GTGCGGGTTAGACCATATTAAGAAAAACAATCAAAACCAAGCGAAATCTGTGATGTG 1067
Db      |||
141  ValArgValArg---TyLySGIuLySProIleGIaAsnGIaIuLySAspVal 159
QY      1068 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCAAGCTCTCAA 1127
Db      |||
160  GIuTyThrValGIuInPheThrProLeuAsnProAspAspPheArgProGIyLeuLyS 179
QY      1128 GATACTAAGCTATTTGAAAAACATAGCTATGGAGACCAATCATCATGATCAAGATTA 1187
Db      |||
180  AspThrLySLeuLeuLySThrLeuAlaIleGIyAspThrIleThrSerGIuInleu 199
QY      1188 GCTCAAGCAACAAGCATTTTAAACAATAACAACCAACCGCTATACGATTTTGAACGTGAC 1247
Db      |||
200  AlaGIuAlaGIuSerIleLeuAsnLySAsnHisProGIyTyThrIleTyArgIuArgAsp 219
QY      1248 TCCTCAATGTCATCTATGACAATGACATTTTCGTACGATTTTACCAATGATCAAGAG 1307
Db      |||
220  SerSerIleValThrHisAspAsnAspIlePheArgThrIleuProMetAspGIuIn 239
QY      1308 TTACTTACCGGTTTAAATATCGGAAACAAGCTTATAGATCAATAAATAATCGGCTG 1367
Db      |||
240  PheThrTyThrValValLySAsnArgGIuInAlaTyArgIleAsnLySlySAspGIyLeu 259
QY      1368 AATGAAGAATAAACAACACTGACTGATCTGTGAGAAATATTAAGCTTAAATAAAGG 1427
Db      |||
260  AsnGIuInIleAsnAsnThrAspLeuIleSerGIuLySTyTYrValleuLySlyGIy 279
QY      1428 GAAAGCCCGTATGATCCCTTATGCGACAGTCACTTGAAACGTGTTACCAATCAATACGTT 1487
Db      |||
280  GIuLySProTyAspProPheAspArgSerHisLeuLySLeuPheThrIleLyTyVal 299
QY      1488 GATGTCGATACCAAGCAATTCCTTAAAGAGAGAGCTTAAACAGCTTACAGCAAGCTAAC 1547
Db      |||
300  AspValAspThrAsnGIuLeuLeuLySArgGIuInleuThrAlaSerGIuArgAsn 319
QY      1548 TTAGACTTCAGAGATTTATACGATCCTCGGATAGAGCTTAAACTCTCAACAATCTC 1607
Db      |||
320  LeuAspPheArgAspLeuTyThrAspProArgAspLySAlaLySLeuLeuTyThrAsnLeu 339

```

OY	1608	GATGCTTTGGTATTATGACTATACCTTA	CTGAAAGTAGAGATAATCA	CGATGAC	1667
Db	340	AspAlaPheGlyIleMetAspTyrThrIleuThrGlyValGluAspAsnHisAspAsp			359
OY	1668	ACCAACCGTATCATACCGTTTATATGGCGACGACCGAGAGAGAAATGCTAGCTAC			1727
Db	360	ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyr			379
OY	1728	CATTAGCT	1736		
Db	380	HisLeuAla	382		

Search completed: January 28, 2006, 02:09:01
 Job time : 167.802 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:55 ; Search time 29.9006 Seconds
(without alignments)
11590.966 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 3726
Sequence: 1 cgaagaccatcatgtctgtc.....ccttcacgatgtcgttag 2096

Scoring table:
BLOSUM62
Xgapop 10 0, Xgapext 0.5
Ygapop 10 0, Ygapext 0.5
Fgapop 6 0, Fgapext 7.0
Delop 6 0, Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame+ n2p.model -DEV=xlD
Q=/cgn2_1/USPTO_spool_p/US09940235/runcat_27012006_144219_27602/app_query.fasta_1.7708
DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -BINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 -ECGN 1 1.159 @runcat_27012006_144219_27602 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database : Issued_Patents_AA:*
- 1: /cgn2_6/prodata/1/iaa/5 COMB.pcp:*
 - 2: /cgn2_6/prodata/1/iaa/6 COMB.pcp:*
 - 3: /cgn2_6/prodata/1/iaa/H COMB.pcp:*
 - 4: /cgn2_6/prodata/1/iaa/PCTUS COMB.pcp:*
 - 5: /cgn2_6/prodata/1/iaa/RE COMB.pcp:*
 - 6: /cgn2_6/prodata/1/iaa/backfile1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1994	53.5	795	2	US-09-211-542A-2
2	1991	53.4	483	1	US-07-854-596B-43
3	1984	53.2	413	2	US-10-360-101-264
4	1984	53.2	414	2	US-09-211-542A-6
5	1984	53.2	414	6	5240845-1
6	1984	53.2	440	1	US-08-560-098A-52
7	1978	53.1	483	1	US-07-854-596B-47
8	1976.5	53.0	1194	1	US-08-488-940-1
9	1969	52.8	499	1	US-07-854-596B-28
10	1966.5	52.8	813	1	US-08-488-940-3
11	1958	52.5	415	1	US-07-854-596B-26
12	1958	52.5	435	1	US-07-854-596B-19

13	1958	52.5	440	1	US-07-854-596B-15	Sequence 15, Appl
14	1958	52.5	859	1	US-07-854-596B-35	Sequence 35, Appl
15	1947.5	52.3	1194	1	US-08-488-940-18	Sequence 18, Appl
16	1946.5	52.2	1194	1	US-08-488-940-17	Sequence 17, Appl
17	1928	51.7	413	1	US-08-759-559-12	Sequence 12, Appl
18	1928	51.7	413	2	US-09-294-457-12	Sequence 12, Appl
19	1928	51.7	413	2	US-09-919-703-12	Sequence 12, Appl
20	1902	51.0	384	2	US-09-374-038-4	Sequence 4, Appl
21	1902	51.0	384	2	US-09-658-179-4	Sequence 4, Appl
22	1898	50.9	372	2	US-09-374-038-3	Sequence 3, Appl
23	1898	50.9	372	2	US-09-658-179-3	Sequence 3, Appl
24	1897	50.9	800	1	US-08-488-940-4	Sequence 4, Appl
25	1897	50.9	1181	1	US-08-488-940-2	Sequence 2, Appl
26	1890.5	50.7	747	1	US-07-854-596B-40	Sequence 40, Appl
27	1887	50.6	401	2	US-09-374-038-1	Sequence 1, Appl
28	1887	50.6	401	2	US-09-658-179-1	Sequence 1, Appl
29	1887	50.6	413	2	US-09-374-038-2	Sequence 2, Appl
30	1887	50.6	413	2	US-09-658-179-2	Sequence 2, Appl
31	1875	50.3	369	1	US-07-854-596B-31	Sequence 31, Appl
32	1698.5	45.6	736	2	US-09-211-542A-4	Sequence 4, Appl
33	1686	45.2	356	2	US-09-211-542A-12	Sequence 12, Appl
34	785	21.1	150	2	US-09-211-542A-14	Sequence 14, Appl
35	640.5	17.2	2231	1	US-08-153-799-16	Sequence 16, Appl
36	640.5	17.2	2386	1	US-10-360-101-235	Sequence 12, Appl
37	638.5	17.1	2355	2	US-10-360-101-235	Sequence 235, App
38	638.5	17.1	2386	2	US-09-961-403-1	Sequence 1, Appl
39	636	17.1	2446	1	US-08-551-356-2	Sequence 2, Appl
40	636	17.1	2446	4	PCT-US93-12687-2	Sequence 2, Appl
41	632	17.0	2324	1	US-08-283-857-1	Sequence 1, Appl
42	632	17.0	2324	4	PCT-US95-09819-1	Sequence 1, Appl
43	627	16.8	2327	6	5455158-1	Patent No. 5455158
44	283	7.6	286	2	US-09-078-091-4	Sequence 4, Appl
45	275	7.4	286	2	US-09-078-091-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 795 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-211-542A-2

Alignment Scores:

Pred. No.:	1-65e-190	Length:	795
Score:	1994.00	Matches:	390
Percent Similarity:	97.75%	Conservative:	1
Best Local Similarity:	97.50%	Mismatches:	5
Query Match:	53.52%	Indels:	4
		Gaps:	2

US-09-940-235-12 (1-2096) x US-09-211-542A-2 (1-795)

```

QY 546 CAGACACATGAGC-----GGATCGGCCCTTCACCGAGTTCGTATGCTGCA 596
Db 366 GlnThrAnsSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
QY 537 CCGAGTGCGCTGCTGACCGCTCATCTGTCACACAGCCAAATGGTTGTTAGCGTTGCT 656
Db 385 ProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuValValSerValAla 404
QY 657 GGTACTGTGAGGGGAGCAATGACATGATTAATTTTGAATCGATCTTAACA 716
Db 405 GlnThrValGluGlyThrAsnGlnAspIleSerLeuLysPheGluIleAspLeuThr 424
QY 717 TCACGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCAAAACCATTT 776
Db 425 SerArgProIleAsnIleGlyGlyThrGluGlnGlyLeuSerProLysSerLysProPhe 444
QY 777 GCTACTGATGATGGCGCGATGTCACATAACTTGAAGAAAGCTGACTTAAGGCTTAT 836
Db 445 AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeuLysAlaIle 464
QY 837 CAAGAACAAATTGATGGCTTAACGTCCACAGTAACGACTACTTGAAGGTCATGATTTT 896
Db 465 GlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspLysPheGluValIleAspPhe 484
QY 897 GCAAGCGATGCAACCATTAAGTGAAGAAAGGCAAGGCTTACTTTCTGACAAAGATGAT 956
Db 485 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheIleAspLysAspGly 504
QY 957 TCGGTAACTTGGCGGACCCCAACCTGTCCAAAGATTTTCTTAAGCGGACATGTGCGCGT 1016
Db 505 SerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHisValArgVal 524
QY 1017 AGACCATTAAGAAAGAAACCATTAACAAACCAAGGAATCTGTGATGTGGAATTAATCT 1076
Db 525 ArgProLysLysGlnLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThr 544
QY 1077 GTACAGTTTACTCCCTTAAACCTGATGACATTTACAGACCAAGGCTTCAAGATATCAAG 1136
Db 545 ValGlnPheThrProLeuAsnProAspAspAspPheAspProGlyLeuLysAspThrLys 564
QY 1137 CTATTGAAACACTAGCTATGCGTGAACACCATCAATCTCAAGAAATTAAGTCAAGCA 1196
Db 565 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeuAlaGlnAla 584
QY 1197 CAAAGCATTTTAAACAAACCAACCAAGGCTATGATTTAAGAGAGTGAATCTCAATC 1256
Db 585 GlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnLysAspSerSerIle 604
QY 1257 GTCACCTGATGACAAATGATTTTCCGTACGATTTTACAAATGATGAAGGTTTACTTAC 1316
Db 605 ValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGlnLysPheThr 624
QY 1317 CGTGTAAATAATCGGAAACAAGCTTATGATCAATAAATAATCGTGTGATGAAGAA 1376
Db 625 ArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGlnGlu 644
QY 1377 ATTAACAACACTGACCTGATCTCGAGAAATATTTACGTCCTTAAAAAGGGAAAAAGCCG 1436

```

```

Db 645 IleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGlyGluAspPro 664
QY 1437 TATGATCCCTTGTGATGCGACGTCACTGAAACGTGTCCACCATGAATAAGTTGATGCGAT 1496
Db 665 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 684
QY 1497 ACCAAGCAATTTGCTAAATAAGTGAAGCGCTTTAAACACTAGCCGAACGTAATTTAGACTTC 1556
Db 685 ThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsnLeuAspPhe 704
QY 1557 AGAAGATTATACGATCTCGTGAATPAAGCTTAACTACTTACATCAACAATCTCGATGCTTTT 1616
Db 705 ArgAspLeuTyrAspProAspGlyAspValAlaLysLeuLeuTyrAsnAsnLeuAspAlaPhe 724
QY 1617 GGTATTATGACATATACCTTAAGCTGAAAGTAGAGATTAATCAAGTACCAACCGCT 1676
Db 725 GlyLeuMetAspTyrThrIleuThrGlnLysValGluAspAsnHisAspAspThrAsnArg 744
QY 1677 ATCATTAACGCTTATATATGCGCAAGCGACCCGAAGAGAGAAATGCTTACATTTAGCT 1736
Db 745 IleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyrHisLeuAla 764

```

RESULT 2

US-07-854-596B-43

Sequence 43, Application US/07854596B
 Patent No. 5434073

GENERAL INFORMATION:

APPLICANT: Dawson, Keith M
 APPLICANT: Hunter, Michael G
 APPLICANT: Czaplowski, Lloyd G
 TITLE OF INVENTION: Proteins and nucleic acids
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESSES:
 ADDRESSER: Dr. John J. McDonnell
 STREET: Ten South Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: IL
 COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,596B
 FILING DATE: 03-JUN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: McDonnell, John J
 REGISTRATION NUMBER: 26,949
 REFERENCE/DOCKET NUMBER: 92,337
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:
 LENGTH: 483 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-854-596B-43

Alignment Scores:

Pred. No.:	2.5e-190	Length:	483
Score:	1991.00 <td>Matches:</td> <td>403</td>	Matches:	403
Percent Similarity:	84.29%	Conservative:	10
Best Local Similarity:	82.24%	Mismatches:	27
Query Match:	53.44%	Indels:	50
DB:	1	Gaps:	9

US-09-940-235-12 (1-2096) x US-07-854-596B-43 (1-483)

QY 294 ACAGATTGATCCCATAGCTGAGAGAGTGTTCATGCTGCTGGGACTTCCTATGTGT 353
|||||
Db 4 ThrApCythrClnuSerGlyGln-----AsnLeuYs-----LeuCyS--- 16

QY 354 CGAGAGAAAGTGGAGAGAGAGCGGATCATCTTGACCTTGAAATAGATGCAC 413
|||||
Db 17 -----GlnGlySerAsn-----ValCySgLyGlnGlnAsnLeuYs--- 28

QY 414 GATCAGACACAGACATCTTATAGATTGAGACACCTGGAGGACCAAGATATACGA 473
|||
Db 29 -----1LeuGlySerAspGlyGlnYsAsn----- 37

QY 474 GGAACCTGCTCCATGATCATCTGACAGAGCAAGCGGAGAGAGTGAAGTGAGAG 533
|||||
Db 38 -----GlnYsVal-----ThrGlyGlnGly----- 44

QY 534 CACACCTCTGTGACAGACCATGAGCGGATCTGGCCCTTACCGATGT----- 584
::: ::|
Db 45 -----ThrProYsProGlnSerH1sAsnAspGlyAspPheGlnGlnProGlnGln 62

QY 585 -----CGTATTCCTGACCTTGAGTGGCTGTGACCGCTCATCTGTCT 626
|||||
Db 63 TyrLeuGln1leGlnGlyAspGlyLeuAlaGlyProGlnThrLeuLeuAspArgProSerVal 82

QY 627 AACACAGCCCAATTGTTGCTTACGTTGCTGCTGCTGCTGAGGAGCAATCAAGACAT 686
|||||
Db 83 AsnAsnSerGlnLeuValValSerValAlaGlyThrValGlnGlnThrAsnGlnAsp1le 102

QY 687 AGCTTTAAATTTTGAATTCGATCTAACATCAAGCATCTGCTGATGAGGAGAAAGACAG 746
|||||
Db 103 SerLeuYsPhePheGln1leAspLeuThrSerArgProAlaHsGlyGlySerThrGln 122

QY 747 CAAGCTTAGTTCAAAATCAAAATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
|||||
Db 123 GlnGlyLeuSerProYsSerYsProPheAlaThrAspSerGlyAlaMetProH1sYs 142

QY 807 CTGGAAGAGCTGACTTACTTAAGGCTATTGAGAACATGATGCTGCTGCTGCTGCTGCT 866
|||||
Db 143 LeuGlnYsAlaAspLeuLeuYsAla1leGlnGlnGlnLeu1leAlaAsnValH1sSer 162

QY 867 AACGACGACTTGTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 926
|||||
Db 163 AsnAspAspTyrPheGlnVal1leAspPheAlaSerAspAlaThr1leThrAspArgAsn 182

QY 927 GCGAAGGTCTACTTGTCTGACAAAGAGTTCGTTACCTTGGCCGACCCAACTGTCCAA 966
|||||
Db 183 GlyYsValTyrPheAlaAspYsAspGlySerValThrLeuProThrGlnProValGln 202

QY 987 GAATTTTGTCTAAGCGGACATGCGCGCTTACGCAATTAATAAAGAAACCAATTAACAAC 1046
|||||
Db 203 GlnPheLeuLeuSerGlyH1sValAlaArgValArgProTyrYsGlnYsPro1leGlnAsn 222

QY 1047 CAAGGAAATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1106
|||||
Db 223 GlnAlaYsSerValAspValGlnYrThrValGlnPheThrProLeuAsnProAspAsp 242

QY 1107 GATTTCAGACAGGCTCTCAAGATACTAAGTATGTAAGAAACATAGCTATCGGTGACCC 1166
|||||
Db 243 AspPheArgProGlyLeuYsAspThrYsLeuLeuYsThrLeuAla1leGlyAspThr 262

QY 1167 ATCATATCTCAAGATTTACTAGTCAAGACAAAGCATTTTAAACAAACCAACCCAGGC 1226
|||||
Db 263 1leThrSerGlnGlnLeuLeuAlaGlnAlaGlnSer1leLeuAsnYsThrH1sProGly 282

QY 1227 TATAGATTATGAAAGTGAAGTCTCAATCGTCAATCGTCAATCGTCAATCGTCAATCGT 1286
|||||
Db 283 TyrThr1leTyrGlnYrAspSerSer1leValThrH1sAspAsnAsp1lePheArgThr 302

QY 1287 ATTTTACCATGATCAAGAGTTTACTTACCGGTAAATGAGGAAACAAAGCTTATGAG 1346
|||||
Db 303 1leLeuProMetAspGlnGlnThrThrYrH1sValYrAsnAspArgGlnAlaTyrGln 322

QY 1347 ATCAATTAATAAATCTGCTCTGAATGAAGAAATTAACAAACATGACCTGATCTTGAGAA 1406
|||||
Db 323 1leAsnYsYsSerGlyLeuAsnGlnGln1leAsnAsnThrAspLeu1leSerGlnYs 342

QY 1407 TATTAAGTCTTAAATAAAGGGAAGAACCGTATGATCCCTTGTATGCGAGTCACTTGAAA 1466
|||||
Db 343 TyrTyrValLeuYsYsGlyGlnYsProTyrAspProPheAspArgSerH1sLeuYs 362

QY 1467 CTGTTCACCATCAATATAGTTGATGTGCATACCAAGAAATTCCTTAAAGTGAAGAGCTC 1526
|||||
Db 363 LeuThrThr1leYsTyrValAspValAsnThrAsnGlnLeuLeuYsSerGlnGlnLeu 382

QY 1527 TTAACAGCTAGCGAAGCTTAACTTGAAGTTCAGAGATTTATGATCGATCCCTGATTAAGGCT 1586
|||||
Db 383 LeuThrAlaSerGlnYrAsnAsnLeuAspAlaPheGly1leMetAspTyrThrLeuThrGlyYs 402

QY 1587 AAATCTCTTAACAATCTCGATGCTTTTGATTTATGACTATACCTTAACCTGAGAAA 1646
|||||
Db 403 YsLeuLeuTyrYrAsnAsnLeuAspAlaPheGly1leMetAspTyrThrLeuThrGlyYs 422

QY 1647 GTAGAGATTAATCAAGATGACACCAACCGTATCAATACCGTTTATATGCGCAAGCACCC 1706
|||||
Db 423 ValGlnAspAsnH1sAspAspThrAsnArg1le1leThrValTyrMetGlyYsArgPro 442

QY 1707 GAAGAGAGAAATGCTAGCTACATTTAGCT 1736
|||||
Db 443 GlnGlyGlnAsnAlaSerTyrH1sLeuAla 452

RESULT 3
US-10-360-101-264
; Sequence 264, Application US/10360101
; Patent No. 6861236
; GENERAL INFORMATION:
; APPLICANT: Mo11, Gertr N.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; PRIOR FILING DATE: 2003-02-07
; PRIORITY FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 413
; TYPE: PRN
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptokinase
US-10-360-101-264

Alignment Scores:
Pred. No.: 1,15e-189 Length: 413
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x US-10-360-101-264 (1-413)

QY 588 ATGCTGAGACCTGAGTGGCTGTGACCGCTGCATCTGTCAACAAACAGCAATGTTGTT 647
|||||
Db 1 1leAlaGlyProGlnThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20

QY 648 AGCGTGTCTGCTGCTGAGGAGCAATCAAGACATTTAGCTTAAATTTTGAATC 707
|||||
Db 21 SerValAlaGlyThrValGlnGlyYrThrAsnGlnAsp1leSerLeuYsPhePheGln1le 40

QY 708 GATCTAACATCAGACCGTGTCAATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
|||||
Db 41 AspLeuThrSerArgProAlaHsGlyGlyYsThrGlnGlnYsLeuSerProYsSer 60

QY 768 AAACATTTGGTACTGATGATGCGGATGTCACATTAATCTTGAAAGGAGTACTACTA 827
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGluLysAlaAspLeu 80
QY 828 AAGGCTATTCAAGAAACAATTGATCGTAAAGTCCACAGTAAACGACATCTTGAAGTC 887
DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspPyrPheGluVal 100
QY 888 ATTGATTTTGGCAGGATGCAACCATTTACTGATGCAAGCGGAGGTCTTACTTGTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValYrPheAlaAsp 120
QY 948 AAAGATGGTGGTAACTTGGCCGACCCCAACCGTCCAGAAATTTTGGTAAAGCGGACAT 1007
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
QY 1008 GTGGCGCTTAAACCATATTAAGAAAAACCATTAACAACCAAGCGAAATCTGTGATGTG 1067
DB 141 ValArgValArgProGlyTrpGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATTAATCTGACATGTTACTCCCTTAAACCTGATGACGATTTTCAGACCGAGTCTCAA 1127
DB 161 GluTrpThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 180
QY 1128 GATCTAAGCTATTGAAACAACAATAGCTATCGGTGACACCATGACATCTCAAGAAATTA 1187
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 200
QY 1188 GCTCAAGCAACAAGCATTTTAAACAACCAACCCAGGCTATACGATTTATGAACGTGAC 1247
DB 201 AlaGlnAlaGlnSerIleLeuLysnLysAsnHisProGlyTrpThrIleTrpGluAsp 220
QY 1248 TCCTCAATGTCATCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1308 TTTTACTACCGTGTAAAAAATGGGGAACAAGCTTATAGATCAATAAAAAATCTGTCTG 1367
DB 241 PheThrTrpArgValLysAsnArgGlnGlnAlaTrpArgIleAsnLysLysSerGlyLeu 260
QY 1368 AATGAAGAAATTAACAACAACAATGATCTCTGAGAAATATTTACGCTCTTAAAAAGG 1427
DB 261 AsnGlnGlnLysAsnAsnThrAspLeuIleSerGlnLysTrpTrpValLeuLysLysGly 280
QY 1428 GAAAGCCGTTATGATCCCTTTGATCGCACTCACTGAAACTGTTCACATCAATACGTT 1487
DB 281 GluLysProGlyTrpAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTrpVal 300
QY 1488 GATGTCGATACCAAGCAATGCTTAAAAAGTGAACGCTTTTAAACGCTGACGATCAAC 1547
DB 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrLysSerGlnGlnGln 320
QY 1548 TTAGACTTCAGAGATTTATAGATCTCGTGAATAAGGCTTAACTACTCAACAATCTC 1607
DB 321 LeuAspPheAspAspLeuTrpAspProAspAspLysAlaLysLeuLeuTrpAsnLeu 340
QY 1608 GATGCTTTGGTATTATGATGATCTTAACTGAAAAAGTAGAGGATTAATCAAGTAC 1667
DB 341 AspAlaPheGlyIleMetAspPyrThrLeuThrGlyValGlnAspAsnHisAspAsp 360
QY 1668 ACCAAGCTATCATACCGTTTATATGGCAAGCGAACCGAAAGAGAGAAATGCTACTAC 1727
DB 361 ThrAsnArgIleIleThrValTrpMetGlyLysArgProGlnGlyGlnLysnAlaSerTrp 380
QY 1728 CATTAGCT 1736
DB 381 HisLeuAla 383

RESULT 4
US-09-211-542A-6
; Sequence 6, Application US/09211542A
; Patent No. 6210667

GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROMBERG & SUNSTEIN, LLP
STREET: 125 Summer Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/211,542A
FILING DATE: 15-December-1998
CLASSIFICATION: 1653
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/069,497
FILING DATE: 15-December-1997
ATTORNEY/AGENT INFORMATION:
NAME: Attorney, Strimpel, Harriet M.
REGISTRATION NUMBER: 37,008
REFERENCE/DOCKET NUMBER: 1874/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 443-9292
TELEFAX: (617) 443-0004
INFORMATION FOR SEO ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-09-211-542A-6
Alignment Scores:
Pred. No.: 1,156-189 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
Gaps: 0
US-09-211-542A-6 (1-2096) x US-09-211-542A-6 (1-414)
QY 588 ATTGCTGACCGTACGAGCGGCTGCTAGACCCGTCATCTGTCAACAACGCAATGCTTGT 647
DB 1 IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 648 AGCGTGTGCTACTGTGAGGGAAGCAATCAAGACATTAAGTCTTAAATTTTGAATC 707
DB 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 40
QY 708 GATTAACATCAAGACCTGCTCATGAGAGAAAGAGAGAGAGGCTTAAGTCCAAAAATCA 767
DB 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlnLeuSerProLysSer 60
QY 768 AAACATTTGCTACTGATAGTGGCGGATGTCACATTAATCTTGAGAAAGCTGACTTACTA 827
DB 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGluLysAlaAspLeu 80
QY 828 AAGGCTATTCAAGAAACAATTGATCGTAAAGTCCACAGTAAACGACATCTTGAAGTC 887
DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspPyrPheGluVal 100
QY 888 ATTGATTTTGGCAGGATGCAACCATTTACTGATGCAAGCGGAGGTCTTACTTGTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValYrPheAlaAsp 120
QY 948 AAAGATGGTGGTAACTTGGCCGACCCCAACCGTCCAGAAATTTTGGTAAAGCGGACAT 1007

|||||
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuSerGlyHis 140
1008 GTGCGCGTTAGACCATATAAGAAAAACAATACAAGAGGAAATCGTTGATGTG 1067
Db 141 ValArgValArgProGlyArgGlyGlyProIleGlnAsnGlnAlaLysSerValAspVal 160
1068 GAATATAGCTGTACAGTTTACTCCCTTAAACCCGTAGACGATTTGACAGCGGTCTAAA 1127
Db 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
1128 GATTACTTAACCTATTGAAAAACCTAGCTATCGGTGACCACTCATCTTCAAGAAATTCTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGGCTATCGATGATTTTGAACCGGAC 1247
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 220
1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
Db 221 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGTCTG 1367
Db 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
1368 AATGAAGAATAAACAACAACGATGACCTGATCTCGAAGAAATATAGCTCTTAAAAAGG 1427
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 280
1428 GAAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACGTTCACCATCAATACGT 1487
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
1488 GATGTCGATACCAACGATTTGCTTAAAGTAGACAGCTCTTAAACAGTACGGAACGTAC 1547
Db 301 AspValAspThrAsnGluLeuLeuLysSerGlnGlnLeuThrLysSerGlnLysArgAsn 320
1548 TTAAGACTTCAAGATTTATACGATCTGTGATAGGCTTAACCTTCAACAATCTC 1607
Db 321 LeuAspPheArgAspLeuLysAspProArgAspLysAlaLysLeuLeuLysAsnLeu 340
1608 GATGCTTTGATATTATGAGCATATACCTTAACTGGAAGAGTAAATGACGATGAC 1667
Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
1668 ACCAACCCTATCATTAACCGTTTATATGGCAAGCGAACCGAAGAGAGATGCTAGCTAC 1727
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyGlnAsnLysSerTyr 380
1728 CATTTAGCT 1736
Db 381 HisLeuAla 383

RESULT 5
5240845-1
; Patent No. 5240845
; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO.; KATANO, TAMIKI;
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
; TSUTOMU
; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS
; NUMBER OF SEQUENCES: 65
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/549,049
; FILING DATE: 06-JUL-1990
; SEQ ID NO:1:
; LENGTH: 414
5240845-1

Alignment Scores:
Pred. No.: 1,156-169 length: 414
Score: 1984.00 Matches: 383

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-12 (1-2096) x 5240845-1 (1-414)

588 ATTGCTGAGCACTGAGTGGCTGTGACCGTCACTGTCATCTGTGCAACAAGCAATGGTTGTT 647
Db 1 IleAlaIleGlyProGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
648 AGCGTTGCTGTACTGTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATC 707
Db 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuLysPhePheGlnIle 40
708 GATTATACATACGACACTGCTGCTCAAGAGAAAGACAGCAAGGCTTAAGTCAAAATCA 767
Db 41 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 60
768 AAACCATTTGCTAGTATAGTGGCGATGTCATATAACTTGAAGAAAGCTGATTAATACTA 827
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 80
828 AAGGCTATTCAAGAACCAATTGATCGTAAAGCTTCAAGTAAAGCACTACTTGAAGGTC 887
Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGlnVal 100
888 ATTGATTTTGGACAGCAATGACCACTTATCTGATGAAACGGCAAGGTCTACTTGTCTGAC 947
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
948 AAGAATGCTTCGTTAAACCTTCCGACCCGACCTGTCCAGAAATTTTGTCTAAGGCAAT 1007
Db 121 LysAspLysSerValThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 140
1008 GTGCGCGTTAGACCATATAAGAAAAACAATACAAGAGGAAATCGTTGATGTG 1067
Db 141 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
1128 GATATCTAAGCTATTGAAAAACCTAGCTATCGGTGACCACTCATCTTCAAGAAATTCTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 200
1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGGCTATCGATGATTTTGAACCGGAC 1247
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 220
1248 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
Db 221 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
1308 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGTCTG 1367
Db 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
1368 AATGAAGAATAAACAACAACGATGACCTGATCTCGAAGAAATATAGCTCTTAAAAAGG 1427
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 280
1428 GAAAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCGTCTTAAACGTTCAATAGCT 1487
Db 281 GluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
1488 GATGTCGATACCAACGATTTGCTTAAAGTAGACAGCTCTTAAACAGTACGGAACGTAC 1547
Db 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrLysSerGlnLysArgAsn 320
1548 TTAAGACTTCAAGATTTATACGATCTGTGATAGGCTTAACCTTCAACAATCTC 1607

Db 321 LeuAspPheArgAspLeuTyraSpProArGAspLysAlaLysLeuLeuTyraAsnLeu 340
Qy 1608 GATGCTTTGGTATTATGAGCTATACCTTAACGTGAAAAGTAAAGTAAATCAGATGAC 1667
Db 341 AspAlaPheGlyIleMetAspTyThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
Qy 1668 ACCAACCGGTATCAATACCGTTTATATGCGAAGCAAGCCGAAAGAGAAATGCTAGCTAC 1727
Db 361 ThrAsnArgIleIleThrValTyrmecGlyLysArgProGluGlyGluAsnAlaSerTy 380
Qy 1728 CATTTAGCT 1736
Db 381 HisLeuAla 383
RESULT 6
US-08-560-098A-52
Sequence 52, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNIENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-52
Alignment Scores:
Pred. No.: 1,19e-189 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
Gaps: 0
US-09-940-235-12 (1-2096) x US-08-560-098A-52 (1-440)
Qy 588 ATTCGTGACCTGATGCTGCTAGACCGTCATCTGTCAACACAGCAATGCTGTT 647
Db 27 IleAlaGlyProGluThrLeuAspAlaGProSerValAsnAsnSerGlnLeuValVal 46

Qy 648 AGCGTTCGTGTACTGTAGAGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATC 707
Db 47 SerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGluIle 66
Qy 708 GATCTAACATCAAGCACTGCTCATGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 767
Db 67 AspLeuThrSerArgProAlaHisGlyGlySerThrGluGlnGlyLeuSerProLysSer 86
Qy 768 AAACCATTTCTACTGATAGTGGCGCATGTCACTAACTTGAGAAAGCTGACTTACTA 827
Db 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 106
Qy 828 AAGGCTATTCAAGAACTTGATCGCTTAAGTCCAGATCAAGCAAGCTCTTTGAGGTC 887
Db 107 LysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrrPheGluVal 126
Qy 888 ATGATTTTTCAGAGCAAGTCAACCATTAAGTCAATGCAAGCAAGTCTTCTTGTGTCAG 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlySerValTyrrPheAlaAsp 146
Qy 948 AAAGATGCTTCGGTAACTTTCGCGAACCACTGTCAGAAATTTTGTAAAGCGACAT 1007
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Qy 1008 GTGCGGCTTAGACCATTAAGAAACCAATCAAAACCAAGGAAATCTGTGATGTCG 1067
Db 167 ValArgValArgProTyrrLysGluLysProIleGlnAsnGlnAlaLysSerValAspVal 186
Qy 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTCAGACAGCTCAAA 1127
Db 187 GluTyrrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 206
Qy 1128 GATCTAAGCTATTGAAACACATAGCTATGGTGACACATCACTCAAGAAATTACTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
Qy 1188 GCTCAAGCACAAAGCATTTTAAACAAACCAAGGCTTATGATTTATGAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrrThrIleTyrrGluArgAsp 246
Qy 1248 TCCGTCAATGCTCACTCAAGCAATGACATTTTCGTCGATTTTCAACAGATCAAGAG 1307
Db 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy 1308 TTACTTACCGGTGTTAAATTCGGGAAACAGCTTATAGATCAATAAATATCGTCTG 1367
Db 267 PheThrTyrrArgValLysAsnAlaArgGluGlnAlaTyrrArgIleAsnLysSerGlyLeu 286
Qy 1368 AATGAAGAATTAACACACATGACCTGATCTCGAAGAAATTAACGTCCTTAAAAAGG 1427
Db 287 AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrrTyrrValLeuLysLysGly 306
Qy 1428 GAAAGGCGGTATGATCCCTTGTATCGGACGTCATGAAACGTTCAACATCAATAGCTT 1487
Db 307 GluLysProTyrrAspProPheAspArgSerHisLysLysLeuPheThrIleLysTyrrVal 326
Qy 1488 GATGCTGATCAACAGCAATGCTTAAAGTAGAGAGCTTTAACAGCTAGCAAGCTAGAC 1547
Db 327 AspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
Qy 1548 TTAGACTTCAGAGATTTATTCGATCTCGTGATTAAGGCTTAACTACTCTACACAAATCTC 1607
Db 347 LeuAspPheArgAspLeuTyrrAspProArgAspLysAlaLysLeuLeuTyrrAsnAsnLeu 366
Qy 1608 GATGCTTTGGTATTATGAGCTATACCTTAACGTGAAAGAGTAATGATTAACAGATGAC 1667
Db 367 AspAlaPheGlyIleMetAspTyrrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
Qy 1668 ACCAACCGGTATCAATACCGTTTATATGCGAAGCAAGCCGAAAGAGAAATGCTAGCTAC 1727
Db 387 ThrAsnArgIleIleThrValTyrmecGlyLysArgProGluGlyGluAsnAlaSerTyrr 406
Qy 1728 CATTTAGCT 1736


```

OY 1947 ATTGAGACACCTGGACAGAGAGATTAATGAGAAACCTGCTCAGTGCATGTCAGACA 2006
Db 448 LeuGlySerAspGlyGlnIuValAsn-----GlnCysVal---Thr 459
OY 2007 GGCAACGGCCGAGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGACCATCGAGC 2066
Db 460 GlyIugIy-----ThrProIyPProGlnSerHis 469
OY 2067 GGATCTGGCCCTTCACCATGTT 2090
Db 470 AsnAspGlyAspPheGlnIuIle 477

RESULT 8
US-08-488-940-1
; Sequence 1, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.,
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Frazer, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-940-1

Alignment Scores:
Pred. No.: 1,17e-188 Length: 1194
Score: 1976.50 Matches: 389
Percent Similarity: 97.50% Conservative: 5
Best Local Similarity: 97.25% Mismatches: 1
Query Match: 53.05% Indels: 5
Db: 1 Gaps: 3

US-09-940-235-12 (1-2096) x US-08-488-940-1 (1-1194)
OY 546 CAGACCAATCGAGAC-----GGATCTGGCCCCCTTCACCCAGATGTCGATTGCTGGA 596
Db 366 GlnThrIasnSerSerSerValProGlyArgGlySerIle---GluGlyArgIleAlaGly 384
OY 597 CCTGAGTGGCGCTGAGACCGTTCATCTGTCAACAACAGCCAAATTTGGTTGAGCGCTGCT 656
Db 385 ProGluTrpIeuLeuAspArgProSerValAsnAsnSerGlnIeuValSerValAla 404
OY 657 GGATCGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTTTGAATGATCTACA 716

```

Dd	405	GlyThrValIguGlyThrAsnGlnAspIleSerIleuysRhePheGluIleAspLeuThr	424
Qy	717	TCAGACCTGCTCATGAGGAAAGACAGCAAGCTTAAAGTCCAAATCAAAACATT	776
Dd	425	SerArgProAlaHisGlyIleGlyThrGluGlnIleuSerProIysSerIysProPhe	444
Qy	777	GCTACTGATAGTGGCGGATGTCATCAATTAACCTTGAAAGCTGACTTAACGGCTATT	836
Dd	445	AlaThrAspSerGlyAlaMetSerHisGlyLeuGluIlyValAlaAspLeuLeuValAlaIle	464
Qy	837	CAAGAACAATGATCGGCTAACGTCCACAGTAAAGACGACTTGTGAGGTCAATTGATTTT	896
Dd	465	GlnGlnGlnIleuIleAlaAsnValHisSerAsnAspAspIyrRheGluValIleLeaspHe	484
Qy	897	GCAAGCGATGCAACCATTAATCTGATCGAAACGGCAAGGTCTACTTTGCTGCAAAAGAGGT	956
Dd	485	AlaSerAspAlaThrIleThrAspArgAsnGlyIySerValIyrRheAlaAspIyAspGly	504
Qy	957	TCGGTAACCTTGGCCGACCCCAACTGTGCCAAGAAATTTTGGTAAAGCGGACATGTGCGCGTT	1016
Dd	505	SerValThrIleuProThrGlnProValGlnGlnPheLeuSerGlyHisValArgVal	524
Qy	1017	AGACCATTAATGAAGAAACCAATCAAAACCAAGCGAAATCTGTGATGTGGAAATACCT	1076
Dd	525	Arg--IyrIySgIyIySproIleGlnAsnGlnAlaIySserValAspValGluIyThr	543
Qy	1077	GTACAGTTACTCCCTTAAACCCGTATGACGATTTTCAGACAGGTCTCAAGATATCAAG	1136
Dd	544	ValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyIleuIySAspThrIyS	563
Qy	1137	CTATTGAAAAACAGTAGCTATTCGATGACACCATCATCTCAAGAAATTAAGTCTCAAGCA	1196
Dd	564	LeuLeuIySThrIleuAlaIleGlyAspThrIleThrSerGlnGlnIleuAlaGlnAla	583
Qy	1197	CAAGGATTTTAAACAAAACCAACCCAGGCTTATACGATTTATGAACGTGACTCTCAATC	1256
Dd	584	GlnSerIleLeuAsnIySAsnHisProGlyIyThrIleIyrgIuAspAspSerSerIle	603
Qy	1257	GTCACTCATGACATGACATTTTCCGACATTTTTCACATGATCAAGAGTTTACTTAC	1316
Dd	604	ValThrHisAspAsnAspIlePheAspIyThrIleLeuProMetAspGlnGlnPheThrIyR	623
Qy	1317	CGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGGTCTGAATGAAGAA	1376
Dd	624	ArgValIySAsnAlaGlnGlnAlaIyArgIleAsnIySserGlyLeuAsnGlnGlu	643
Qy	1377	ATTAACAACACTGACCTGATTTCTGAGAAATATTACGTCTTATAAAAAGGGGAAAGCCG	1436
Dd	644	IleAsnAsnThrAspLeuIleSerGlnIyIyIyIyValLeuIySlySgIyIySPro	663
Qy	1437	TATAGATCCCTTGAATCGAGTCACTTGAAACGTGTCAACATCAAAATCGTTGATGTGCAT	1496
Dd	664	TyrAspProPheAspArgSerHisIleuIySLeuPheThrIleIySValAlaAspValAsp	683
Qy	1497	ACCAACGAATTCGTAAAAAGTGACAGACTTTTAAACAGCTTACGCAACGTAACTTAAGCTTC	1556
Dd	684	ThrAsnGlnIleuLeuIySserGlnGlnIleuThrAlaSerGluArgAsnLeuAspPhe	703
Qy	1557	AGAGATTTATTCGATCTCTCGTATAGGCTTAAACTACTTACACAACTTCGATGCTTTT	1616
Dd	704	ArgAspLeuIyTyrAspProArgAspIyAlaIySLeuIyIyTyrAsnAsnLeuAspAlaPhe	723
Qy	1617	GGTATTATGACATTAACCTTAAACGTGAAAAAGTGAGGATTAATACAGATGACACCAACCGT	1676
Dd	724	GlyIleIleAspIyThrIleuThrGlyIyValGlnAspAsnHisAspAspThrAspArg	743
Qy	1677	ATCATTAACCGTTATATATGCGCAAGCGACCCGAAAGAGAGATGCTACTTCAATTAAGCT	1736
Dd	744	IleIleThrValIyMetGlyIySArgProGlnGlyGlnAsnHisAspIyThrHisLeuAla	763

Sequence 28, Application US/07854596B
Patent No. 5434073
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Czaplowski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-596B-28

Alignment Scores:
Pred. No.: 4,066-188 Length: 499
Score: 1969.00 Matches: 382
Percent Similarity: 96.00% Conservative: 2
Best Local Similarity: 95.50% Mismatches: 16
Query Match: 52.84% Indels: 0
Gaps: 0

US-09-940-235-12 (1-2096) x US-07-854-596B-28 (1-499)

QY	537	ACCTCTGTCAGACACATCGAGCGGATCTGCCCTTACCGATGTCGATTTGCTGA	596
Db	69	Thrlleaaserllealalalaysglugluglyvalserleuaspysargllealagly	88
QY	597	CCTGAGGAGCTGCTAGACCGCTCATCTGTCAACAAGCAATGTTGTTACGTTGCT	656
Db	89	ProgluttrpleuleuaspargproserValaasnserGlnleuValalSerValaA	108
QY	657	GTAAGTGTGAGGGAGCAATCAAGACATTAAGTCTTAATTTTGAATCGATTAACA	716
Db	109	GlythrValgluglyThrshnlnaspIleSerleuylpPheheGlnleuaspLeuThr	128
QY	717	TCAGACCTGCTCATGAGAGAAAGACAGAGAGGCTTAAGTCAAAATCAAAACGATT	776
Db	129	SerargProalaHleGlyGlyserThrGlnGlnGlyLeuSerProlyserIysPhe	148
QY	777	GCTACTGATAGTGGCGCATGTCACATTAACCTTGAAGAAAGCTGACTTAAAGCTATT	836
Db	149	AlathraspSerGlyAlaMetProHlsylsleuGlnulysAlaaspLeuIysAlaIle	168
QY	837	CAAGACAATTGATGCTTAAGTCCACAGTAACGAGACTACTTGAAGCTATGATTTT	896
Db	169	GlnGlnGlnleuIleAlaasnValHlsSerAsnaspIysPheGlnValIleaspPhe	188

QY	897	GCAAGCATGCAACCATTAAGTATGATGAAAGCGCAAGGTCTACTTTGTCGACAAAGTGT	956
Db	189	AlaSeraspAlatrHleThrAspArgasnGlylsValIysPheAlaaspIysaspGly	208
QY	957	TCGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTAAAGGACATGTGCCGCTT	1016
Db	209	SerValThrleuProThrGlnProValGlnGlnPheleuSerGlyHleValArgVal	228
QY	1017	AGACCATTTAAAGAAAAACCATACAAACCAAGCAATCTGTTGATGTGGAATATACT	1076
Db	229	ArgProIyrlsGlnlysProIleGlnasnGlnAlaIysSerValaspValGlyIyrls	248
QY	1077	GTACAGTTACTCCCTTAAACCTGATACGATTCACACAGGCTCAAGATCTAG	1136
Db	249	ValGlnPheThrProleuasnProaspaspPheargProGlyleuIysaspThrIys	268
QY	1137	CTATTGAAAACATAGCTATGCTGACACCATCATCTCAAGATTTACTAGCTCAAGCA	1196
Db	269	leuIystrHleuAlaIleGlyAspThrIleThrserGlnGlnleuIleuAlaGlnIa	288
QY	1197	CAAGCATTTTAAACAAAACACCCAGGCTATACATTTATGAACGTGACTCTCAATC	1256
Db	289	GlnserIleuasnIysThrHisProGlyIyrlsThrIleIyrlsGlnArgaspSerIle	308
QY	1257	GTCACATGACATGACATTTCCGTACGATTTTCAACATGATCAAGATTTACTTAC	1316
Db	309	ValThrHisaspAsnaspIlePheargThrIleuProMetaspGlnIlePheThrIy	328
QY	1317	CGTGTAAATAACGGGAAACAAGCTTATAGATCATATAAATATCGTCTAAAGGGAAGCG	1376
Db	329	HleValIysasnaspGlnGlnAlaIyrlsGlnleuIysIysSerGlyleuIasnGlnIu	348
QY	1377	ATAAACAACCTGACCTGATCTCTGAGAAATATTACGCTTAAAGGGAAGCG	1436
Db	349	IleasnThrAspIleuIleSerGlnIyryIyrlsValleuIyryGlyIyryPro	368
QY	1437	TATGATCCCTTTGATCCGACGTCATTTGAAATCTGTCCACCAATATAGCTGATGCTGAT	1496
Db	369	TyraspProPheaspArgSerHlsleuIyryleuPheThrIleIyryValaspValasn	388
QY	1497	ACCAACAATTTGCTAAAGAGAGACAGCTCTTAACAGCTACGGAACGTAATTTAGACTTC	1556
Db	389	ThrshnIleuIyrySerGlnGlnleuIyryThrIleuIyryValaspIyryPhe	408
QY	1557	AGAGATTATACGATCTCGTATAGGCTTAACTACTTCAACAATCTGATGCTTTT	1616
Db	409	ArgaspIyryrAspProalaspIyryAlaIyryleuIyryrAsnaspIyryPhe	428
QY	1617	GCTATTATGACATTAACCTTAAGTGAAGAAAGTAAAGATTAATACGATGACCAACCGT	1676
Db	429	GlyIleuIyryrThrIleuThrGlyIyryValGlnIyryrAsnaspIyryrAspThrAspArg	448
QY	1677	ATCATTAACGTTTATATAGGCAAGCGACCGCAAGAGAAATGCTAGTACATTTAGCT	1736
Db	449	IleIleThrValIyryMetGlyIyryArgProGlnGlnIyryIleuAlaSerIyryHleuAla	468

RESULT 10
US-08-488-940-3
Sequence 3, Application US/08488940
Patent No. 5854049
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:


```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,940
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Frazer, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 813 amino acids
TYPE: amino acid
STRANDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-940-3

Alignment Scores:
Pred. No.: 9,466-168 length: 813
Score: 1966.50 Matches: 382
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 52.78% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-12 (1-2096) x US-08-488-940-3 (1-813)
QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCTCCATCTGTGCAACACAGCAATTGTTGTT 647
DB 1 IleaaglyproglutripheuleuaspargproserValasnsbserleuVal 20
QY 648 AGCGTGTGCTGCTGCTGAGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 707
DB 21 SerValaaglythrValgluglythraanginaepileserleuysphegile 40
QY 708 GATCTAACATCAGCAGCTGCTCATGTGAGAAAGACGACGAGCTTAAGTCCAAATCA 767
DB 41 AspleutrSerargproalahlsglyglysthrnglgllyleuSerPolySer 60
QY 768 AAACCATTTGCTAGTATGAGCGCGAGTGTCACTAAACCTTGAAGAGCTGACTACTA 827
DB 61 LysprophelaThrAspSerGlyalMetSerHisbyleuGlylyalaaspleu 80
QY 828 AAGCTATTCAAGAACCAATTGATGCTTAAAGTCCACAGTAAAGCACTACTTGAAGTC 887
DB 81 LysalaileglnlgluGlnleuilealaenValHisSerAsnaspaspTyrphegluVal 100
QY 888 ATTGATTTTGAAGGAGGACCAACATTAAGTCCAAAGCGAAGGTCTACTTGTGAC 947
DB 101 IleaspPhealaSerAspAlaThrIleThrAspArgasnGlylyValTyrPhealaasp 120
QY 948 AAAGATGTTGCGGTAACTTGGCGGACCAACCTGCTCAAGAAATTTTGTGAAGCGGACAT 1007
DB 121 LysaspGlySerValThrleuProthrnglnproValGlnGlnPheuleuSerGlyHis 140
QY 1008 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAAATCTGTTGATGTG 1067
DB 141 ValargValarg---TyrlyseGlylyspProlleglnasnGlnAlaIalysSerValaspVal 159
QY 1068 GAATTAATCTGACAGTTTACTCCCTTAAACCTGATGACGATTTCAGCCAGGTCTGAAA 1127
DB 160 GluTyrThrValaGlnPheThrProleuasnProaspAspAspPheargProGlyleuLys 179
QY 1128 GATCTAGCATTTGAAAAACATAGCTATCGGTGACACATCAACATCTCAAGAAATCACTA 1187

```

```

DB 180 AspThrLysleuLeuLysThrleuAlaileGlyAspThrIleThrSerGlnGluLeuLeu 199
QY 1188 GCTCAAGCAACAAGCAATTTTAAACAACACCCAGGCTATACGATTTATGACAGTAC 1247
DB 200 AlaGlnAlaGlnSerIleleuasnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 219
QY 1248 TCCTCAATCGTCACTGACATGACAAATGACATTTTCGTCAGATTTTACCAAGATCAAGAG 1307
DB 220 SerSerIleValThrHisAspAsnAspIlePheargThrIleleuProMetAspGlnGlu 239
QY 1308 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATGATCAATAAATAATCTGTCCTG 1367
DB 240 PheThrTyrArgValIalysAsnArgGlnGlnAlaIalTyrArgIleasnLysLysSerGlyLeu 259
QY 1368 AATGAAGAAATTAACAACACTGACCTGATCTCGAAGAAATATACGTCCTTAAATAAGG 1427
DB 260 AsnGlnGlnIleasnThrAspLeuIleSerGlnLysTyrTyrValIleuLysLysGly 279
QY 1428 GAAAAGCCGTATGATCCCTTGTGATCGGACGCTGAAACTGTTCACCATCAATACGTT 1487
DB 280 GluLysProTyrAspProPheAspArgSerHisleuLysleuPheThrIleLysTyrVal 299
QY 1488 GATGTCGATACCAAGCAATGTCTAAAGTGAACAGCTCTTAAACGCTTACGGAACGTAC 1547
DB 300 AspValAspThrAsnGlnleuLeuLysSerGlnGlnleuThrAlaSerGluArgAsn 319
QY 1548 TTAGACTTCGACGATTTATACGATCTCTCGGATTAAGGCTTAACCTCTACCAATCTC 1607
DB 320 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysleuLeuTyrAsnAsnleu 339
QY 1608 GATGCTTTGGTATTTATGAGCTATACCTTACTGGAAGAGTAAAGATATCAAGATGAC 1667
DB 340 AspAlaPheGlyIleMetAspTyrThrleuThrGlyLysValGluAspAsnHisAspAsp 359
QY 1668 ACCAACCGTATCAATAACCGTTTATATGCGCAACGACCCCAAGAGAGAAATGCTAGTAC 1727
DB 360 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnGluAsnAlaSerTyr 379
QY 1728 CATTAGCT 1736
DB 380 HisleuAla 382

RESULT 11
US-07-854-596B-26
; Sequence 26, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,337
TELECOMMUNICATION INFORMATION:

```


Percent Similarity: 98.96% Conservative: 1
 Best Local Similarity: 98.69% Mismatches: 4
 Query Match: 52.55% Indels: 0
 DB: 1 Gaps: 0

US-09-940-235-12 (1-2096) x US-07-854-596B-19 (1-435)

```

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGCTCATCTGTCAACAAGCAAGCAATGTTGTT 647
DB 22 IleaaglyProgluttrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 41
QY 648 AGCGTGTGCTACTGTTGAGGGAAGCAATCAAGCATTAGTCTTAATTTTGAATC 707
DB 42 SerValaIaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuysPheGlnIle 61
QY 708 GATCTAACATCCAGACTCTCTCATGAGGAAAGACAGACAGCGCTTAAGTCCAAATCA 767
DB 62 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 81
QY 768 AAACCATTTGCTACTGATGATGCGCGAGTGTCACTTAACCTTGAGAAAGCTGACTTA 827
DB 82 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 101
QY 828 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAGACGACTACTTGAGTGC 887
DB 102 LysAlaIleGlnGlnGlnLeuLeuIleAlaAsnValHisSerAsnAspAspLysrPheGlnVal 121
QY 888 ATTGATTTTGCAGGAGTCAACCATTAATCTGAAACGAGGAGTCTACTTGTGTGAC 947
DB 122 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValLysrPheAlaAsp 141
QY 948 AAAAGATGTTGGTGAACCTTGGCGAACCCGACCTGTCCAAAGATTTTGTGAAGCGACAT 1007
DB 142 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 161
QY 1008 GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGCAATCTGTGATGTC 1067
DB 162 ValaGlnValArgProGlyLysGlnLysProIleGlnAsnGlnAlaLysSerValaAspVal 181
QY 1068 GAATTAAGTGTACAGTTTATCTCCCTTAAACCTTGATGACGATTTTGACCAAGCTCTCAA 1127
DB 182 GlnLysrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeuLys 201
QY 1128 GATCTAAGCTATTGAAAAACATAGTATCGGTGACACATCAATCTCAAGAAATTACTA 1187
DB 202 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 221
QY 1188 GCTCAAGACAAAGCATTTTAAACAAAACCAACCCGAGGTATACGATTTATGAAGCTGAC 1247
DB 222 AlaGlnAlaGlnSerIleLeuAsnLysThrHisAspGlnLysrThrIleLysrGlnAlaAsp 241
QY 1248 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTTCACATGATCAAGAG 1307
DB 242 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 261
QY 1308 TTTACTTACCGTGTAAATAATGGGAAACAGCTTATGATCAATTAATAAATGTGCTCG 1367
DB 262 PheThrLysHisValaLysAsnArgGlnGlnAlaLysrGlnIleAsnLysLysSerGlyLeu 281
QY 1368 AATGAAGAAATTAACAACAACATGACCTGATCTGTGAGAAATATTAGCTCTTAAAAAGG 1427
DB 282 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysrLysrValLeuLysLysGly 301
QY 1428 GAAAAAGCGTATGATCCCTTTGATGCGAGTCACTTGAACCTGTTCCACATCAATACGTT 1487
DB 302 GlnLysrProLysrAspProPheAspArgSerHisLysLeuLysPheThrIleLysrLysrVal 321
QY 1488 GATGTCGATTAACAAGAAATTTGCTAAAAAGTAGACAGCTCTTAAACAGCTACGGAAGTAA 1547
DB 322 AspValaAsnThrAsnGlnLysLeuLysSerGlnGlnLeuLeuThrAlaSerGlnAlaAsn 341
QY 1548 TTAGACTTCAGAGATTTATACGATCTCGTGAATAGGCTTAAGCTTAATCAACATATTC 1607

```

DB 342 LeuAspPheArgAspLeuLysrAspProArgAspLysAlaLysLeuLeuLysrAsnAsnLeu 361

QY 1608 GATGCTTTGGTATTTATGACCTATTAACCTGGAACATGAGGATTAATCCAGTAC 1667

DB 362 AspAlaPheGlyLysLeuMetAspLysrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 381

QY 1668 ACCAACCGTATCAATACCGTTTATATGGGCAAGCAACCCGAAGAGAGAAATGCTAGCTAC 1727

DB 382 ThrAsnArgIleIleThrValLysrMetGlyLysArgProGlnGlnAlaSerLysr 401

QY 1728 CATTAGCT 1736

DB 402 HisLeuAla 404

RESULT 13

US-07-854-596B-15

Sequence 15, Application US/07854596B

Patent No. 5434073

GENERAL INFORMATION:

APPLICANT: Dawson, Keith M

APPLICANT: Hunter, Michael G

APPLICANT: Czaplowski, Lloyd G

TITLE OF INVENTION: Proteins and nucleic acids

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:

ADDRESSER: Dr. John J. McDonnell

STREET: Ten South Wacker Drive, Suite 3000

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/854,596B

FILING DATE: 03-JUN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McDonnell, John J

REGISTRATION NUMBER: 26,949

REFERENCE/DOCKET NUMBER: 92,337

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ. ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 440 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-854-596B-15

Alignment Scores:

Pred. No.: 4,79e-187

Score: 1958.00

Percent Similarity: 98.96%

Best Local Similarity: 98.69%

Query Match: 52.55%

DB: 1

US-09-940-235-12 (1-2096) x US-07-854-596B-15 (1-440)

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGCTCATCTGTCAACAAGCAAGCAATGTTGTT 647

DB 27 IleaaglyProgluttrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46

QY 648 AGCGTGTGCTACTGTTGAGGGAAGCAATCAAGCATTAGTCTTAATTTTGAATC 707

DB 47 SerValaIaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuysPheGlnIle 66


```

Db      122  ILeaSPheAlaSeRAspAlaThrIleThraSPaRgAsnGlyLeValTyRPhelAlaSP 141
Qy      948  AAAGATGGTTCGGTAACCTTGCCGACCCAACTGTCGAAGAAATTTTGGTAAAGCGGACAT 1007
Db      142  LysaSPgIySeRValThrLeuProThrgInProValGInGluPhenLeuSeRGIyHIs 161
Qy      1008  GTGGCGGTAGACATATTAAGAAAAACAATACAAAACAAGCGCAATCTGTGATGAG 1067
Db      162  ValaRgValaSProTyRLeuSProlIeGInaSPnGInAlaYsSeRValaSPVal 181
Qy      1068  GAATTAAGTGTACAGTTTACTCCCTTAAACCTGATGACATTTTACAGACAGGTCTCAA 1127
Db      182  GluTyRThraValGInPhethrProLeuaSPnProaSPaSPaSPPhethrGProGlyLeuYs 201
Qy      1128  GATTAAGCTATTTGAAAACACTAGCTATGGTGACACATCCATCTCAAGATTAATA 1187
Db      202  AspThrLysLeuLeuTherThreulaIleGlyAspThrIleThSeRGIInGluLeu 221
Qy      1188  GCTCAAGACAAAGATTTTAAACAACCAAGCGGTATACGATTAATTAAGACGTGAC 1247
Db      222  AlaGInAlaGInSeRtleuLeuSPnThraSProGlyTyRThrIleTyRGIuaRgAsp 241
Qy      1248  TCCCTCAATCGTCACTCATGACATGACATTTTCCGTAAGATTTTACCAATGATCAAG 1307
Db      242  SeSerIleValThrHIsaSPaSPnSPriePhethrGThrIleuProkethaSPGIInGlu 261
Qy      1308  TTTACTTACCGTGTAAATCGGGAACAAGCTTATAGATCATATAAAATCTGCTG 1367
Db      262  PheThrTyRHisValaYsaSPnRgGInGInAlaTyRGIInIleSPnLysSeRGIyLeu 281
Qy      1368  AATGAAGAAATTAACAACAACATGACCTGATCTCTGAGAAATTTAGCTCTTAAAAAGG 1427
Db      282  AsnGluGluIleAsnaSPnThraSPnLeuSeRGIuTyRtyRValleuYsGIy 301
Qy      1428  GAAAAGCCGTATGATCCCTTATGAGCACTGCTGAACCTGTTCACATCAATACGTT 1487
Db      302  GluTySProTyRAsPProPhethaSPaRgSeRHisLeuThraSPnThrIleuSPtyRVal 321
Qy      1488  GATGTCGATACCAAGAAATGCTTAAAGAGCAAGCTTAAACAGCTACGCAACCTTAAC 1547
Db      322  AsPValaSPnThraSPnGluLeuLeuLysSeRGIInGluLeuThraSPnSeRGIuARgAsn 341
Qy      1548  TTAAGCTTCAGAGATTTTATGATCTCTGATTAAGCTTAACTACTTACCAACATCTTC 1607
Db      342  LeuSPPhethaSPnLeuTyRAsPProaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSP 361
Qy      1608  GATGCTTTGGATATGAGCATATACCTTAAGTGAAGAAAGTAAAGATATACAGATGAC 1667
Db      362  AsPAlaSPnGlyIleMetAsPtyRThrLeuThraSPnGlyValaGluAsPnHIsaSPaSP 381
Qy      1668  ACCAACCGTATCATTAACCGTTATATAGGCAAGCAACCCGAAGAGAGATGCTAGCTAC 1727
Db      382  ThrAsnaSPGIleIleThraValTyRmetGlySaRgProGluGlyGluSPnHIsaSPeTyR 401
Qy      1728  CATTTAGCT 1736
Db      402  HisLeuAla 404

```

RESULT 15
US-08-488-940-18
Sequence 18, Application US/08488940

Patent No. 5854049
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P. C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA

```

? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent in Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/488, 940
? FILING DATE: 09-JUN-1995
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Fraser, Janis K.
? REGISTRATION NUMBER: 34,819
? REFERENCE/DOCKET NUMBER: 05433/009001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617/542-5070
? TELEFAX: 617/542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 18:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1194 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-488-940-18

```

Alignment Scores:

```

Pred. No.: 9,34e-186 Length: 1194
Score: 1947.50 Matches: 397
Percent Similarity: 87.12% Conservative: 9
Best Local Similarity: 85.19% Mismatches: 50
Query Match: 52.27% Indels: 10
DB: 1 Gaps: 5

```

US-09-940-235-12 (1-2096) x US-08-488-940-18 (1-1194)

```

Qy      546  CAGACCAATTCGAGC-----GGATTCGGCCCTTACCCGATGTTGATTTGCTGGA 536
Db      366  GInThraSPnSeRSeRSeRValProGlyARgGlySeRtle---GluGIyARgIleAlaGly 384
Qy      597  CCTGATGGCTGTATACCGTCCATCTGTCGAACAGCAAGCAATGGTGTAGCGGTCT 656
Db      385  ProGluTrPheLeuSPnAlaSProSeRValaSPnSeRtleuValaSPeRVala 404
Qy      657  GGTACTGTTGAGGGGAGCAATCAAGCATTAAGCTTAAATTTTGAATCGATCTAACA 716
Db      405  GlyThraValGluGlyThraSPnGInaSPrieSeRleuAlaPhethrGluIleAsPLeuThr 424
Qy      717  TCAAGACCTGCTCATGAGGAAAGACAGACAGAGCTTAAGTCCAAATCAAAACATTT 776
Db      425  SeRAlaSProAlaHIsGlyGlyAlaThraGluGInGlyLeuSeRProAlaSeRtySProPh 444
Qy      777  GCTACGATAGTGGCGGATGTCACATTAACCTTGAAGACGTGACTTAAGGCTATT 836
Db      445  AlaThraSPSeRtyAlaMetSeRHisLysLeuGluYsaSPnLeuLysAlaIle 464
Qy      837  CAAGAACAAATTGATCCGTAACGTCAACAGTAAAGCACTTGAAGTCAATGATTTT 896
Db      465  GInGluGInLeuIleAlaSPnValHIsSeRAsnaSPaSPtyRThrheGluValIleAsPhe 484
Qy      897  GCAAGCGATGCAACCAATTAAGTCAAAACGCAAGGTCTTACTTGTGAACAAAGATGT 956
Db      485  AlaSeRAsPAlaThrIleThraSPaRgAsnGlyLeValTyRPhelAlaSPlySaSPGly 504
Qy      957  TCGGTACCTTGCCGACCCAACTGTCCAAAGATTTTGTAGCGGACATGCGCGGT 1016
Db      505  SeRValThrLeuProThrgInProValGInGluPhenLeuSeRGIyHIsaSPValaSPVal 524
Qy      1017  AGACCATATTAAGAAAAACAATACAAAACAAGCGCAATCTGTGATGTAATATACT 1076
Db      525  Arg---TyRLeuGluSProIleGInaSPnGInAlaYsSeRValaSPValaSPGlyTyRThr 543

```

```
OY 1077 GTACGATTACTCCCTTAAACCTGATGACGATTTCAGACCAAGGCTCAAGAATACTAAG 1136
|||
Db 544 ValGlnPheThrProLeuSnpProAspAspPheArgProGlyLeuLysAspThrLys 563
|||
OY 1137 CTATTGAATAAACCCTACTACCTCGGTGACACCATGACATCTCAAGAACTTACTAGCTCAAGCA 1196
|||
Db 564 LeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeuAlaGlnAla 583
|||
OY 1197 CAAGCATTTTAAACAATAACCAACCAAGGCTATACGATTATGAAAGTGAATCTCCGAATC 1256
|||
Db 584 GlnSerIleLeuSnpLysSnpHisProGlyThrIleLeuThrGlnArgAspSerSerIle 603
|||
OY 1257 GTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGAATCAAGATTACTTAC 1316
|||
Db 604 ValThrHisAspSnpAspIlePheArgThrIleLeuProMetAspGlnGlnPheThrTyr 623
|||
OY 1317 CGTGTAAATAATCGGGAACAAGCTTATAGATCAATTAATAATCTGGTCTGAATGAAGAA 1376
|||
Db 624 ArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysSerGlyLeuAsnGlnGln 643
|||
OY 1377 ATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGGGAAAGCCG 1436
|||
Db 644 IleAsnAsnThrAspLeuLieserGlnLysTyrTyrValLeuLysGlnLysPro 663
|||
OY 1437 TATGATCCCTTGTATGACAGTCACTTGAACGTGTCAACATCAATACGTTGATGTGAT 1496
|||
Db 664 TyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAsp 683
|||
OY 1497 ACCAACAATTTGCTAAAAAGTGAAGCACTTTAAACAGCTAGCGAAGCGTAACCTTAGACTTC 1556
|||
Db 684 ThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsnLeuAspPhe 703
|||
OY 1557 AGAGATTTTATAGATCTCTGCTGTAAGGCTAAACTACTACCAACAATCTCGATGCTTTT 1616
|||
Db 704 ArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuLysAsnAsnLeuAspAlaPhe 723
|||
OY 1617 GGTATTATGACATATCTTAACTGGAATAAGTAGAGATATATCAGATGACCAACCGT 1676
|||
Db 724 GlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAspThrAsnArg 743
|||
OY 1677 ATCATTAACCGTTTATATGGGCAAGCGAAGCGAAGAGAGATCTAGCTACCAATTAAGCT 1736
|||
Db 744 IleIleThrValTyrMetGlyLysArgProGlnGlnLysAsnAlaSerTyrHisLeuAla 763
|||
OY 1737 GGTGGTGGCAGCGCGAACAAGATTGATCCATAGCTGAGAAGCTTTTGATCATGCTGCT 1796
|||
Db 764 -----TyrAspAlaAspArgTyrThrGlnGlnGlnArgGlnValTyrSerTyrLeuArg 781
|||
OY 1797 GGACCTTCTATGTGTGTCGAGAAACGTGGAGAGAGCCCTACCAAGGCTGATGATGTA 1856
|||
Db 782 TyrThrGlyThrProIleProAspAsnProAsnAspLysAsnAsnSerGlnLeuValVal 801
|||
OY 1857 GATTGT-----ACTTGCTGGAGAGAGCGAGCGACATCACTTGCACTTCAAGAAAT 1910
|||
Db 802 SerValAlaGlyThrValGlnGlyThrAsnGln-AspIleSerLeuLysPhePheGlnI 821
|||
OY 1911 AGATGCAACGATCAGG 1926
|||
Db 821 eAspLeuThrSerArg 826
|||
```

Search completed: January 28, 2006, 02:42:08
Job time : 57.9006 secs


```

QY 648 AGCGTGTGCTGTAAGGGGAGCAATTAAGTCTTAATTTTGAATC 707
  |||
Db 21 SerValAlaGlyThrValAlaGlyThrAsnGlnAspLysSerLeuysPhePheGlyIle 40
QY 708 GATCTAATGATCAAGCACTGCTGATGAGGAAAGACAGCAAGCTTAAGTCCAAATGCA 767
  |||
Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnIleuSerProLysSer 60
QY 768 AAACCATTTGCTACTGATAGTGAGCGCGGATGTCATTAATCTTGAGAACTGACTTA 827
  |||
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGlnIleuLysAlaAspLeuLeu 80
QY 828 AAGGCTATTTCAGAACAAATTTGCTACGTCACAGTAACAGCAAGCTTAAGTGTGTC 887
  |||
Db 81 LysAlaAlaIleGlnIleGlnIleuLeuAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100
QY 888 ATTGATTTTGCAGGATGCAACCATTAAGTCAAGCAAGCTTAAGTGTGTCAG 947
  |||
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 948 AAAGATGTTGCTGTAACCTTGCCGACCAACCTGTCAGAAATTTTGTAAAGCGGACAT 1007
  |||
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnIlePheLeuSerGlyHis 140
QY 1008 GTGCGCGCTTAAGCACTATTAAGAAAAACCAATACAAACCAAGCAAGCTTGTATGTC 1067
  |||
Db 141 ValArgValArgProGlyLysGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY 1068 GAATATCTGTACAGTTTACTCCCTTAACCTGATGACGATTTTGACAGGCTCCAA 1127
  |||
Db 161 GlnTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY 1128 GATATGAAGCTATTGAAAACTAGCTATCGGTGACACCATGACATCTCAAGATTACTA 1187
  |||
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnIleuLeu 200
QY 1188 GCTCAGACCAAGCAATTTTAAACAAAACCAAGCTTATGCAATTTATGAAGCTGAC 1247
  |||
Db 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnAsp 220
QY 1248 TCCCTCAATCGTCACTGATGACAAATGCTTCCGATGATTTTACATGATGCAAGG 1307
  |||
Db 221 SerSerIleValThrHisAspAsnAspIlePheAlaGlyThrIleuProLeuSerGlnGln 240
QY 1308 TTTTACTTACCGTGTAAAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCGGTCTG 1367
  |||
Db 241 PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 1368 AATGAGAAATTAACACACATGACCTGATCTGTGAGAAATATTACGCTTTAAAAAGG 1427
  |||
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrValIleuLysLysGly 280
QY 1428 GAAAGCGGTATGATCCCTTGAATGCGACATCTGAAACCTGTCCCAATCAATCGTT 1487
  |||
Db 281 GlnLysProLysArgProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
QY 1488 GATGTCATTAACCAAGAAATTTGCTAAAAAGTGAAGCTTAAACAGCTACCGCAATAC 1547
  |||
Db 301 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnIleuLeuThrAlaSerGlnAsnGln 320
QY 1548 TTAGACTTACAGATTTTATAGACTCTCGTATTAAGGCTAAACTACTTACCAACTGTC 1607
  |||
Db 321 LeuAspPheArgAspLeuLysArgProArgAspLysAlaLysLeuLeuLysArgAsnLeu 340
QY 1608 GATGCTTTTGTGATTAATGACTATTAAGCTTAAGTGAAGGATTAATCAAGATGAC 1667
  |||
Db 341 AspAlaAspGlyLysIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 360
QY 1668 ACCAAGCTATATCAACCGTTATATGCGCAAGCGACCGCAAGAGAGATGCTACTGAC 1727
  |||
Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnIleuLysAsnAlaSerTyr 380

```

```

QY 1728 CATTAAGT 1736
  |||
Db 381 HisLeuAla 383

RESULT 2
US-09-940-235-2
; Sequence 2, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamnara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-09-940-235-2

Alignment Scores:
Pred. No.: 2,44e-172 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-12 (1-2096) x US-09-940-235-2 (1-414)
QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCCAAATGTTGTT 647
  |||
Db 1 IleAlaGlyProGlnIlePheLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 648 AGCGTGTGCTGTAAGGGGAGCAATTAAGTCTTAATTTTGAATC 707
  |||
Db 21 SerValAlaGlyThrValAlaGlyThrAsnGlnAspLysSerLeuysPhePheGlyIle 40
QY 708 GATCTAATGATCAAGCACTGCTGATGAGGAAAGACAGCAAGCTTAAGTCCAAATGCA 767
  |||
Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnIleuSerProLysSer 60
QY 768 AAACCATTTGCTACTGATAGTGAGCGCGGATGTCATTAATCTTGAGAACTGACTTA 827
  |||
Db 61 LysProPheAlaThrAspSerGlyAlaMetSerHisLeuGlnIleuLysAlaAspLeuLeu 80
QY 828 AAGGCTATTTCAGAACAAATTTGCTACGTCACAGTAACAGCAAGCTTAAGTGTGTC 887
  |||
Db 81 LysAlaAlaIleGlnIleGlnIleuLeuAlaAsnValHisSerAsnAspAspTyrPheGlnVal 100
QY 888 ATTGATTTTGCAGGATGCAACCATTAAGTCAAGCAAGCTTAAAGCTTACTTGTGTCAG 947
  |||
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 948 AAAGATGTTGCTGTAACCTTGCCGACCAACCTGTCAGAAATTTTGTAAAGCGGACAT 1007
  |||
Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnIleuLysAsnAlaSerTyr 140

```

```

QY 1008 GTGCGGCTTGACCATTAATAAGAAAAACAATACAAACCAAGGAAATCTGTGATGTG 1067
DB 141 ValArgValArgProTyrLysGluYsProIleGlnAenglnAlaLysSerValAspVal 160
QY 1068 GATATATAGTACAGTTTATCCCTTAACCCGTGATGACATTTTCAACACCGGTCTCAA 1127
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGluYleuYs 180
QY 1128 GATATAGCTATTTGAAAACATAGTATGCTGATGACACCATCATCATCTCAAGAAATTA 1187
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1188 GCTCAAGCACAAACATTTTAAACAAAACCAACAGGCTATACATTTATGAAAGTGAC 1247
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGluYrThrIleTyrGluArgAsp 220
QY 1248 TCCCTCAATGCTCATATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 221 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
QY 1308 TTATCTTACCGTGTATTAATAATCGGGAACAGCTTATGATGATCAATAAAAAATCGTCTG 1367
DB 241 PheThrTyrArgValAlaLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeu 260
QY 1368 AATGAAGAAATAAACAACATGACCTGATCTTGAGAAATATTAAGTCTTAAAAAAAGG 1427
DB 261 AsnGluGluIleAsnThrAspLeuIleSerGluYsTyrTyrValIleuLysLysGly 280
QY 1428 GAAACACCGTATGATCCCTTTGATCCGATGCTTGAACSTGTTACACATCAATACGTT 1487
DB 281 GluLysProTyrLysProPheAspArgSerHisLysLysLeuPheThrIleYsTyrVal 300
QY 1488 GATGTGATATCAACGAATTTGCTAAAGAGGACAGCTCTTAACAGCTAAGGAAAGCTAAC 1547
DB 301 AspValAspThrAsnGluLeuLeuLysSerGlnGluLeuThrAlaSerGluArgAsn 320
QY 1548 TTAGACTTGAGAGATTTATACGATCTCGATAGGCTAAACTCTCAACAACTCTC 1607
DB 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
QY 1608 GATGCTTTGGTATTTATGACATATACCTTAACGTGAAGAGTAGAGATTAATCAAGATGAC 1667
DB 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAspAsp 360
QY 1668 ACCAACCGTATGATACCGTTTATATAGGACACGACCCGAAGAGAGATGCTAGCTAC 1727
DB 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 380
QY 1728 CATTAGCT 1736
DB 381 HisLeuAla 383

```

```

; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252

Alignment Scores:
Pred. No.: 5,866-170 Length: 414
Score: 1958.00 Matches: 378
Percent Similarity: 98.96% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 4
Query Match: 52.55% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-12 (1-2096) x US-10-300-215-252 (1-414)
QY 588 ATTGCTGAGCTGAGTGCTGCTGACCGTCCATCTGTCACAAACAGCAATGCTGTT 647
DB 1 IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
QY 648 AGCGTTGCTGATCTGTTGAGGAGGACGAATCAAGCATTTAGTCTTAAATTTTGAATC 707
DB 21 SerValAlaGlyThrValGluGlyThrAsnLysPheSerLeuLysPhePheGluLe 40
QY 708 GATTTAATCATGACACATCTGCTCATGAGAGAAAGACAGCAAGGCTTAAGTCCAAATCA 767
DB 41 AspLeuThrSerArgProAlaHisGlyLysThrGluGlnGlyLeuSerProLysSer 60
QY 768 AAACATTTGCTACTGATAGTGGGCGCATGTCACATTAACCTTGAGAAAGCTGACTACTA 827
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeuLeu 80
QY 828 AAGCTATTCAAGAAACATTTAGCTGATACCTGCACAGTACAGCACTTATGAGGTC 887
DB 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 100
QY 888 ATTGATTTGCAACGATGCAACCATTAATGATGAAACGCAAGGCTTACTTTGCTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
QY 948 AAAAGATGTTGGTAACTTCCGACCAACCTGTCCAAGAAATTTTGCTAAGCGGACAT 1007
DB 121 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
QY 1008 GTGCGGCTTGACCATTAATAAGAAAAACAATACAAACCAAGGAAATCTGTGATGTG 1067
DB 141 ValArgValArgProTyrLysGluYsProIleGlnAenglnAlaLysSerValAspVal 160
QY 1068 GATATATAGTACAGTTTATCCCTTAACCCGTGATGACATTTTCAACACCGGTCTCAA 1127
DB 161 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGluYleuYs 180
QY 1128 GATATAGCTATTTGAAAACATAGTATGCTGATGACACCATCATCATCTCAAGAAATTA 1187
DB 181 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 200
QY 1188 GCTCAAGCACAAACATTTTAAACAAAACCAACAGGCTATACATTTATGAAAGTGAC 1247
DB 201 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGluYrThrIleTyrGluArgAsp 220
QY 1248 TCCCTCAATGCTCATATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
DB 221 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240

```

QY	1308	TTTACTTAACGGTTAAAAATCGGGAACAACCTTATAGACATCAATAAATAATCTGCTCG	1367
Db	241	PhenrIrrYrhIsvalLYsAsnArgGInGlnAlaTyrgIuIleAsnLYsSerGlyIueu	260
QY	1368	AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG	1427
Db	261	AenGInGInuIleAsnAsnThrAspneuIleSerGInuYsTyrgYrValIeuLYsGely	280
QY	1428	GAAGAACCGGTATGATCCCTTTGATGCGACGTCACTTGAACCTGTCAACATTAATAGCTT	1487
Db	281	GIuYsPProTYrAspProPheAspArgSerHisLeuYsIeuPhenrIleLYsTyrgAl	300
QY	1488	GATGTCGATCAACAAGATGTCTAAAAAGTAGACAGCTCTTAACAGCTAGGGAAGTAC	1547
Db	301	AspValAsnThrAsnGInuLeuLeuLYsSerGInuIleuLeuThrAlaSerGInuArgAsn	320
QY	1548	TTAGACTTCAGAGATTATATACGATCCTCTGTATAGAGCTTAAGCTTAACATCTC	1607
Db	321	LeuAspPheArgAspLeuTyrgAspProArgAspLYsAlaLYsIeuLeuTyrgAsnAsnIeu	340
QY	1608	GATGCTTTGGTATTTATGACCTAATACCTTAATCGAAAAAGTAGAGGATTAACAAGATGAC	1667
Db	341	AspAlaAspGelyIleuLeuAspTyrgThrIeuThrGlyLYsValGInuAsnAsnHisAspAsp	360
QY	1668	ACCAACCGTATCATTAACCGTTATATAGGCAAGCGAACCCGGAAGAGAGAATGCTAGCTAC	1727
Db	361	ThrasnArgIleIleThrValTYrMeGelyLYsArgProGInuGInuIleAsnAlaSerTYr	380
QY	1728	CATTTAGCT 1736	
Db	381	HisIeuAla 383	

RESULT 4

```

US-10-988-943-1
: Sequence 1, Application US/10988943
: Publication No. US20050176085A1
: GENERAL INFORMATION:
: APPLICANT: Center for Genetic Engineering and Biotechnology
: TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
: TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
: TITLE OF INVENTION: COMPLEX MIXTURES.
: FILE REFERENCE: Proteomics CU2003-269
: CURRENT APPLICATION NUMBER: US/10/988,943
: CURRENT FILING DATE: 2004-11-15
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: patentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 415
: TYPE: PRT
: ORGANISM: Streptococcus equisimilis
: US-10-988-943-1

```

Alignment Scores:

Pred. No.:	5,866-170	length:	415
Score:	1958.00	Matches:	378
Percent Similarity:	98.96%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	4
Query Match:	52.55%	Indels:	0
BB:	5	Gaps:	0

US-09-940-235-12 (1-2096) x US-10-988-943-1 (1-415)

QY	588	ATTGTGGAACTCGAGTGGCTGCTGAACCGTCCATCTGTCAACAACGCCAATTGGTTGTT	647
Db	2	lIeaIaGIyProGIuTrIpeuIeuSpaRproSerValaIaSpaSerGIuIeuVal	21
QY	648	AGCGTTGCTGACCTGTGAGGGGACGATCAAGACATTAAGTTTAATTTTGGAAATC	707
Db	22	SerValaIaGIyThrValGIuGIyThraSngIaSpIleSerIeuIySphepGluIle	41
QY	708	GATCAACATCAACGACCTGCTCATGAGGAAGACAGACCAAGCGCTTAAGTCCAAATCA	767
Db	42	AspIeuTrIserAArgproAlaIaSIyGIyIySThGIuGIuGIyIeuSerProIySer	61

QY	768	AAACATTGGTACTGATAGTAGGGGCGATGTCACATTAACCTTGGAGAAAGCTTACTA	822
Db	62	LysProPheAlaAtmAspSerGlyAlaMetCProHslyuSeuGulUlySAlaAspLeuLeu	81
QY	828	AAGCCTATTCAAGAAACAATTGATGCTTAACGCTCCACAGTAAACGACGCTACTTTGAAGTC	887
Db	82	LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerSerAsnAspAspTyrPheGlnVal	101
QY	888	ATTGATTTTGGCAGCGATGCAACCATTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC	947
Db	102	IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyuValTyrPheAlaAsp	121
QY	948	AAAGATGGTCCGGTAACCTGGCGAACCCAGCTGTCGAAAGATTTTGGTAAGGGGACAT	1007
Db	122	LysAspGlySerValThrLeuProThrGlnProValGlnGlnLeuLeuSerGlyHis	141
QY	1008	GTGGCGCTTAACCATTAATAAGAAAAACAATAACAAACCAAGCGAAATCGTTGATGATG	1067
Db	142	ValArgValAlaGProTyrGlyGlnUlyProIleGlnAsnGlnAlaUlySerValAspVal	161
QY	1068	GAATATATCTGTACAGTTTACTCTCCTTAAACCTGATGACGATTTCAACACAGGCTCAAA	1127
Db	162	GluTyrThrValAlaGlnPheThrProLeuAsnProAspAspPheAspProGlyLeuLys	181
QY	1128	GATACTAAGCATTTGAAAAACATPACGTAATGGGTGACACCATCATCTCAAGAAATTACTA	1187
Db	182	AspHnTrpLysLeuLysThrLeuAlaIleGlyAspHnTrIethrSerGlnGlnLeuLeu	201
QY	1188	GCTCAAGACACAAGACATTTTAAACAAAACCAACCCAGCGTATACGATTTATGAACGTGAC	1247
Db	202	AlaGlnAlaGlnSerIleLeuAsnUlyThrHisProGlyTyrThrIleTyrGlnUlyArgAsp	221
QY	1248	TCCCTCAATCGTCATCATGACATGACATTTTCCGTCGATTTTACCAATAGGATCAAGAG	1307
Db	222	SerSerIleValThrHisAspAsnAspIleAspHnTrIethrLeuProMetAspGlnGln	241
QY	1308	TTTACTACCGTGTTAAAAAATCGGAGCAAGGCTATATAGATCAATAAAAAATCGGCTG	1367
Db	242	PheHnTrpIleValUlySAsnArgGlnGlnAlaTyrGlnUlyLeuLysUlySerGlyLeu	261
QY	1368	AATGAAAGAAATTAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG	1427
Db	262	AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnUlySerTyrTyrValLeuUlySlyGly	281
QY	1428	GAAGAGCCGTAATGATCCCTTGGATCGGACATCTTGAACATGCTTACCATCAATATAGCTT	1487
Db	282	GluUlySerProTyrAspProPheAspArgSerHisIleuUlySerLeuPheThrIleUlyTyrVal	301
QY	1488	GATGTCAGTACCAAGCAATGCTTAAAAAGTAGAGAGCGCTTTAACAGCTAGCGAAGCGTAAC	1547
Db	302	AspValAsnHnTrpAsnGlnUlyLeuUlySerGlnGlnLeuUlyThrAlaSerGlnUlyArgAsn	321
QY	1548	TTAGACTTCAGAGATTTATAGACTCGTCGTGATTAAGGCTTAACTACTTCAACAACATCTC	1607
Db	322	LeuAspPheArgAspLeuTyrAspProArgAspUlyAlaUlySerLeuLeuTyrAsnAsnLeu	341
QY	1608	GATGCTTTTGGTATTATAGACTATACCTTAACTGGAAAGTAGAGGATTAATCCAGATGAC	1667
Db	342	AspAlaPheGlyIleMetAspTyrThrLeuThrIlyUlyValGlnAspAsnHisAspAsp	361
QY	1668	ACCAACCGTATCATTAACGTTTATATAGGCAAGGCAACCCGGAAGAGAGAAATGCTAGCTAC	1727
Db	362	ThrAsnArgIleIleThrValTyrMetGlyUlyAsnArgProGlnGlnUlyGlnAsnAlaSerTyr	381
QY	1728	CATTAGACT 1736	
Db	382	HisLeuAla 384	

```

; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified strep protein
US-10-300-215-253

Alignment Scores:
Pred. No.: 5,95e-169 Length: 414
Score: 1947.00 Matches: 376
Percent Similarity: 98.69% Conservative: 2
Best Local Similarity: 98.17% Mismatches: 5
Query Match: 52.25% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-12 (1-2096) x US-10-300-215-253 (1-414)
QY 588 ATTGCTGAGCACTGAGTGGCTGCTGAGACCTGTCATCTGTCAACAACCACTGTTGTT 647
DB 1 IleaIaGlyProGluTrpLeuLeuAaPArgProSeValaAaAaSerGlnLeuVala 20
QY 648 AGCGTTGCTGACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 707
DB 21 ServaIaIaGlyThyValaIaGlnIyThraAaGlnAaPrlIeSerLeuYsPhePheGluIle 40
QY 708 GATCTAACATCAAGCACTGCTGATGAGAGAAAGAGAGAGCAAGCTTAACTCCAAATCA 767
DB 41 AspleuThraSerArgProAlaHisGlyGlyLeuThraGlnIleGlnIleuSerProIySer 60
QY 768 AAACCACTTGTCTACTGATAGTGGCGCGGATGTCACTAATTAAGAAAGCTGACTTAA 827
DB 61 LysProPheAlaThrAspSerGlyAlaMetProHisLeuLeuGluLysAlaAspLeuLeu 80
QY 828 AAGGCTATTCAAGAACATTTGATCGTAACTGTCACAGTAAAGCACTTCTTGAAGTC 887
DB 81 LysAlaLysGlnGlnGlnLeuLeuIleAlaAaValHisSerAaAaAaPArgPheGluVal 100
QY 888 ATTGATTTTGAAGGATCAACCATTAATGATGAAACAGGCAAGGCTTACTTTGCTGAC 947
DB 101 IleAspPheAlaSerAspAlaThrIleThraAspArgAaGlnIySValIyTrpPheAlaAsp 120
QY 948 AAAGATGTTGGGTAACTTGTCCGAGCAACCACTGTCCAAAGATTTTGTCTAACGGACAT 1007
DB 121 LysAspGlySerValIleThraProThraGlnProValGlnGlnIleuLeuSerGlyHis 140
QY 1008 GTGCCCGTTAGACATATTAAGAAAAACAATACAAACCAAGCAAAATCTGTGATGTC 1067

```

```

DB 141 ValArgValaArgProTyTrpGlyLeuSerProIleGlnAaGlnAlaLysSerValaAspVal 160
QY 1068 GAATATCTGATGAGCTTTATCTCCCTTAACCTGATGACGATTTAGACAGCTTCAAA 1127
DB 161 GluTyTrpValaGlnPheThraProLeuAaAaPArgAaPheArgProGlyLeuYs 180
QY 1128 GATACCTAAGCTATTGAAAACATGACTGATCGGTGACACATCACTCAAGAAATCTACTA 1187
DB 181 AspThrLysLeuLeuYsThraLeuAlaIleGlyAspThrIleThraSerGlnIleuLeu 200
QY 1188 GCTCAAGCAACAAGCATTTTAAACAAAACCAACCAAGGCTATACGATTTATGACGTGAC 1247
DB 201 AlaGlnIaGlnSerIleLeuAaLysThraHisProGlyTyTrpThraIleTyGluArgAsp 220
QY 1248 TCCCTCAATCGTCACTCATGACAAAGACATTTTCCGTAGATTTTACATGATGATCAAG 1307
DB 221 SerSerIleValaIleThraAspAaAaPrlIePheArgThrIleLeuProMetAspGlnGlu 240
QY 1308 TTACTTACCGGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCGTCTG 1367
DB 241 PheThrTyTrpHisValLysAaAaArgGlnAlaIleTyArgIleAaLysLysSerGlyLeu 260
QY 1368 AATGAAGAAATTAACAACACTGACCTGATCTGTGAGAAATATTAAGTCTTAAAAAAGG 1427
DB 261 AsnGluGluIleAaAaAaThraAspLeuIleSerGlnIyTyTrpValaLeuLysGly 280
QY 1428 GAAAAGCCGTATGATCCCTTTGATCGACATGACATTTGAAACTGTCACCATCAATACGTT 1487
DB 281 GluLysProTyTrpAspProPheAspArgSerHisLeuLysLysPheThrIleLysPheVal 300
QY 1488 GATCGATACCAACGAATTCCTTAAAGAGAGAGCTTTAAACGCTTAAACGCTAGGAACTAC 1547
DB 301 AspValaAaThraAaGlnLeuLeuLysSerGlnIleuLeuThraLysSerGluArgAaAa 320
QY 1548 TTAGACTTCAGAGATTATATGATCTCTGATTAAGGTAAGCTTAACTGTAACAATCTC 1607
DB 321 LeuAspPheArgAspLeuTyTrpAspProArgAspLysAlaLysLeuLeuTyAspAaAa 340
QY 1608 GATGCTTTGGTATTAATGACATATACCTTAACTGAAAGATGAGATTAACAGATGAC 1667
DB 341 AspAlaPheGlyIleMetAspTyTrpThraThrGlyLysValaGluAspAaAaHisAspAsp 360
QY 1668 ACCAAGCGTATCATTAACCTTTTATATGGGCAAGGAGCCGAAGAGAGATGCTAGTAC 1727
DB 361 ThrAspArgIleIleThraValTyMetGlyLysArgProGluGluAaAlaSerTyx 380
QY 1728 CATTAGCT 1736
DB 381 HisLeuAla 383

RESULT 6
US-09-919-703-12
; Sequence 12, Application US/09919703
; Patent No. US20020165129A1
; GENERAL INFORMATION:
; APPLICANT: KRYSTAL, Gerald
; APPLICANT: RABKIN, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413

```

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-12

Alignment Scores:

Pred. No.:	3,27e-167	Length:	413
Score:	1928.00	Matches:	377
Percent Similarity:	98.95%	Conservative:	1
Best Local Similarity:	98.63%	Mismatches:	2
Query Match:	51.74%	Indels:	2
DB:	3	Gaps:	2

US-09-940-235-12 (1-2096) x US-09-919-703-12 (1-413)

```

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 647
Db 1 IleaIaGIyProGIuTtPleuEusArpArProSeVaIaSnAnSeRGIuEuaIVal 20
QY 648 AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGATTAAGCTTAAATTTTGAATC 707
Db 21 SeVaIaIaGIyThrVal--GIyInraSnGIaAspIleSeRleuYsPheGluIle 39
QY 708 GATCTAATCAACGACCTGCTCATGAGGAGAAAGACAGCAAGCTTAAGTCCAAATCA 767
Db 40 AspleuThrSeRArGIProIaIaHISGIyGIyThrGIuGIuInGIyLeuSeRProIySeR 59
QY 768 AAACATTTGCTACTGATAGTGGCGGAGATGTCAATAACTTGAGAAAGCTGACTTA 827
Db 60 LysProhealathrAspSerGIyAlaMetSerHISlySeuGIuIySaIaAspLeu 79
QY 828 AAGGCTATTCAAGAACATTTGCTGCTACGTCACAGTACAGCAAGCTTCTTGAGTGC 887
Db 80 LysaIaIeGIuGIuInGIuEuaIeIaIaSnValHISeRasnaSAspIyRPhGIuVal 99
QY 888 ATTGATTTTGAAGCGATCAACCATTAATCTGATCGAAAGCGCAAGCTCTTGGCTGAC 947
Db 100 IleAspPheIaSeRAspIaIaThrIleInraSAspAsnGIySaIyThrPheIaAsp 119
QY 948 AAAGATGGTTCGGTAACTTGGCGAACCCAACTGTCCAAAGATTTTGGTAAAGCGACAT 1007
Db 120 LysAspGIySeRValThrIeupProThrGIuInProValGIuGIuPheIeueuSeRGIyHIS 139
QY 1008 GTGGCGCTTGAACCATTAATAAGAAAAACCAATACAAACAGCGAAATGTGTATGTC 1067
Db 140 ValaGyValaIarGIProIyGIySglIySProIleGIaSnGIaIaIaYSeRValaSpVal 159
QY 1068 GAATATACGTACAGTTTACCTCCCTTAACCCGTGATGAGATTTTGACACCGGTCGAA 1127
Db 160 GIuIyThrValGIuPheThrProIeuaSnProIaSpAspRPhaIarGIyLeuIyS 179
QY 1128 GATATTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCAATCTCAAGATTAATA 1187
Db 180 LeuThrIySleuIeueuThrIeuaIaIeGIyAspThrIleThSeRGIuGIuEuaIe 199
QY 1188 GCTCAGACCAAGCAATTTTAAACAAAAACACCCAGGCTTAACGATTTATGAAGTGCAC 1247
Db 200 AlaGIaIaGIaInSeRleIeueuSnIySaSnHISProGIyThrIleIyRGIuIaGAsp 219
QY 1248 TCCCTAATCGTCACTGATGACAAATGATTTCCGTAACGATTTTACCAATGATGCAAGAG 1307
Db 220 SerSeRleIeuaIThHISAspAsnAspIlePheAIGThrIleuProMetArGIuGIu 239
QY 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTAATAGATCAATAAAAATCTGGCTGC 1367
Db 240 PheThrIyRarGIaValIySaSnArGIuGIuInaIaIyRarGIleuSnIySSeRGIyLeu 259
QY 1368 AATGAAGAAATTAACAACACTGACCTGATCTGTGAGAAATAT--TACGCTCTTAAAAAA 1424
Db 260 AsnGIuGIuIleAsnAsnThrAspLeuIleSeRleuGIuIyThrIySaIyValIeueuIyS 279
QY 1425 GGGGAAAAGCGTATGATCCCTTTGATGCGACATCTGTAAGAACTGTTCACATCAATAC 1484

```

```

Db 280 GIyIuIySProIyRAspProPheAspAspSeRHISleuIySleuPheThrIleIySTyr 299
QY 1485 GTTATGTCCTATACCAACGAATTTGCTAAAGTAGACGCTCTTAACAGCTAGCGAAGCT 1544
Db 300 ValaSpValaIaSpThraSnGIuEuaIeueuYsSeRGIuGIuInIeueuThraIaSeRGIuIarG 319
QY 1545 AACTTAAGCTTCAGAGATTTTATACGATCTCGTGTAAAGGCTAAACTACTGACAAACAT 1604
Db 320 AsnIeuaSnPheAAspLeuIyRAspProIaGAspIySaIaIySleuIyRAsnAn 339
QY 1605 CTCGATGCTTTGGTATTATGACTATACTTAACGTGAGAAAGTAGAGATTAATACGAT 1664
Db 340 LeuAspIaRPhGIyIleMetAspIyThrIeuthrGIySValGIuAspAsnHISAsp 359
QY 1665 GACACCAACCGTATCATACCTGTTATATGGGACAGCAAGCCGAAGAGAGAAATGTCGC 1724
Db 360 AspThraSnArGIleIleThrValIyMetGIySaRGIuGIuGIuAsnAlaSeR 379
QY 1725 TACCAT 1730
Db 380 TyrHIS 381

```

RESULT 7

US-10-474-792-658

; Sequence 658, Application US/10474792
; Publication No. US20040236072A1

; GENERAL INFORMATION:

; APPLICANT: Olmsted, Stephen

; APPLICANT: Zagursky, Robert

; APPLICANT: Nickbarg, Elliot

; APPLICANT: Winter, Louie

; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES

; FILE REFERENCE: AM 100399

; CURRENT APPLICATION NUMBER: US/10/474,792

; CURRENT FILING DATE: 2003-10-14

; NUMBER OF SEQ ID NOS: 674

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 658

; LENGTH: 440

; TYPE: PRT

; ORGANISM: Streptococcus pyogenes

US-10-474-792-658

Alignment Scores:	Pred. No.:	6,22e-155	Length:	440
Score:	1794.00	Matches:	341	
Percent Similarity:	93.99%	Conservative:	19	
Best Local Similarity:	89.03%	Mismatches:	23	
Query Match:	48.15%	Indels:	0	
DB:	5	Gaps:	0	

US-09-940-235-12 (1-2096) x US-10-474-792-658 (1-440)

```

QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATGGTTGTT 647
Db 27 IleaIaGIyRGIyThrIeupProIaSpArGIProIleAsnAnSeRGIuEuaIVal 46
QY 648 AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGATTAAGCTTAAATTTTGAATC 707
Db 47 SerMetAaIaGIyIleValGIuGIuGIyThraSpIySValaPheIleAsnRPhGIuIle 66
QY 708 GATCTAATCAACGACCTGCTCATGAGGAGAAAGACAGCAAGCTTAAGTCCAAATCA 767
Db 67 AspleuThrSeRGIuInProIaIaHISGIyGIyThrGIuGIuInGIyLeuSeRProIySeR 86
QY 768 AAACATTTGCTACTGATAGTGGGCGATGCAATAACTTGAGAAAGCTGACTTAATA 827
Db 87 LysProhealathrAspAsnGIyAlaMetProHISlySeuGIuIySaIaAspLeu 106
QY 828 AAGGCTATTCAAGAACATTTGATGCTTAACGTCAACAGTAAAGCACTTATGAGGTC 887
Db 107 LysaIaIeGIuGIuInGIuEuaIeIaIaSnValHISeRasnaSAspIyThrPheGIuVal 126

```

```

OY 888 ATTGAATTTTGCAGCGATGCAACCAATTACTGATCGAAACGGCAGGCTTACTTTGCTGAC 947
Db 127 ILeaSPheAlaSerSPheAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 146
OY 948 AAAAGATGCTCGGTAACTTCCGCGACCCAACTGTGTCAGAAATTTTGTCAAGCGACAT 1007
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuGlyGlnHis 166
OY 1008 GTGCGCGTTAGACCATATTAAGAAAAACAATAACAACCAACGGAATCTGTGATGTC 1067
Db 167 ValAlaGValAlaArgProTyrGlyGlySerProValGlnAsnGlnAlaLysSerValAspVal 186
OY 1068 GAATTAAGTGTACAGTTTAACTCCCTTAAACCTGATGACGATTTTCAAGACGGCTCAAA 1127
Db 187 GlnTyrThrValGlnPheThrProLeuAsnProAspAspPheAspPheArgProGlyLeuLys 206
OY 1128 GATACTAAGCTATTGAAAACAAGTACTGCGGTGACACCATGACATCTCAAGAAATTAATA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
OY 1188 GCTCAGACGACAAAGCAATTTTAAACAACCAACCGGCTATCGATTTTATGAAAGCTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisArgGlyTyrThrIleTyrGlnArgAsp 246
OY 1248 TCCCTCAATCGTCACTCATGACAAATGACATTTTCCGTAACGATTTTCAATGATGACAG 1307
Db 247 SerSerIleValThrHisAspAsnAspIlePheAspThrIleLeuProMetAspGlnGln 266
OY 1308 TTTACTTAAACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCGTCTG 1367
Db 267 PheThrTyrHisValLysAsnArgGlnGlnAlaTyrGlnIleAsnProLysThrGlnIle 286
OY 1368 AATGAAGAAATTAACAACATGACCTGATCTGAGAAATATTTAGCTCTTAAAGGAGG 1427
Db 287 LysGlnLysThrAsnAsnThrAspLeuValSerGlnLysTyrTyrValLeuLysGlnGly 306
OY 1428 GAAAGCGGTATGATCCCTTTGATCGAGTCACTTGAACCTGTCACATCAAAATCGTT 1487
Db 307 GlnLysProTyrAspProPheAspArgSerHisLysLysLeuPheThrIleLysTyrVal 326
OY 1488 GATGTGATACCAAGCAATTTGCTAAAGAGTGAAGCACTTTAAACAGTACGGAACCTAAC 1547
Db 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuThrIleAspGlnAlaGln 346
OY 1548 TTAAGCTTCAGAGATTTATACGATCTCGTATAGGCTTAACTTCAACCAATCTTC 1607
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
OY 1608 GATGCTTTTGTATATGACATATACCTTAACTGGAAGAGTAAAGATATCAACGATGAC 1667
Db 367 AspAlaAspLeuAspIleMetAspTyrThrLeuThrGlyLysValGlnLysAsnHisAspLys 386
OY 1668 ACCAACCGTATATACCGTTTATATGCGCAAGCGACCGCAAGAGAGAGATGCTAGCTAC 1727
Db 387 AsnAsnAlaGValValThrValTyrMetGlyAspArgProLysGlyAlaLysGlySerTyr 406
OY 1728 CATTTAGCT 1736
Db 407 HisLeuAla 409

```

```

; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 235
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of fibronectin
US-10-360-101-235

Alignment Scores:
Pred. No.: 7,45e-49 Length: 2355
Score: 638.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
DB: Gaps: 28

US-09-940-235-12 (1-2096) x US-10-360-101-235 (1-2355)

OY 5 GACCATTCATGTTGTTGCTCAGGTCGACAGCTTTTGACAGACGATCGCTTACAGT--- 61
Db 91 AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 110
OY 62 -----TCGTCGCGGTATCGGATTCATTCGTCGTAACGATTAAGGCAACC 106
Db 111 HisGlnGlyGlnGlnSerTyr-LysIleGlyAspThrTrp-----ArgArgPr 126
OY 107 CCGCAGCTTACGCGGCTCTCAACGACAGGACGACATCATGCGCACCCGTGCGCAGA 166
Db 126 OhHisGlnThr----- 129
OY 167 CCGAAGCTGCGCGAGATTCGATCCGCGAATTAATAGCATCTATAGGAGACCA 226
Db 129 ----- 129
OY 227 CAAGGTTTCCCTAGAA---ATAATTTGTTTAACTTAAAGAGATATACCATGG 283
Db 130 -GlyGlyTyrMetLeuGlnCysValCysLeuGlyAsnGlyLysGln-TyrThr--- 147
OY 284 TGCAAGACAAACAGATTGTACCACTAGCTGAGAGTGTGATCATGCTGCTGGAGCTT 343
Db 148 -----CysLysProIleAlaGlnLysCysPheAspHisAlaIleGlyHis 163
OY 344 CCTATGTGTCGAGAAACGTGGA----- 368
Db 163 ertYrValValGlyGlnThrTrpGlnLysProTyrGlnGlyTyrMetMetValAspCysT 183
OY 369 -----GAAGCAGCGAGACGATCATCTTGCATTTAGAAATAGATGACAGATC 417
Db 183 hrCysLeuGlyGlnGlySerGlyArgIleThrCysThrIleSerArgAsnArgCysAsnAsp 203
OY 418 AGGACCAAGACATCTTATAGAAATTTGAGACACCTGAGCAAGAAGATATATGAGAA 477
Db 203 LnsPthrArgThrSerTyrArgIleLysAspThrTrpSerLysLysAspAsnArgGly 223
OY 478 ACCGTGTCAGTGCATTCGACAGGCAACGCGCAGAGAGAGTGAAGTGAAGAGCACA 537
Db 223 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyLysTyrLysCysGlnArgHisT 243
OY 538 CTTGTGTGACAGACCAATCGAGCGGATGTGGCCCTTACCGATGTTCTGATTGCT--- 593
Db 243 hrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaIaValT 263
OY 593 ----- 593
Db 263 yrglnProGlnProHisArgProGlnProProTyrGlnHisCysValThrAspSerGlyV 283
OY 594 -----GGACCTGAGTGGCTGACAGCTGCATCTGTCACAAACAGCCAT 639
Db 283 alValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGlnM 300

```

```

RESULT 8
US-10-360-101-235
; Sequence 235, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly) peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8

```

```
QY 640 TGGTTGATGGCTTGGTCTGTAAGTGTGAGGGGCAATCAAGATTAAGTCTTAATTT 699
Db 300 etleuCy5ThrCysleuGly----- 306
QY 700 TTGAATTCGATCTAATCAATCAAGCTGCTCATGAGGAAAGAGAGCAAGGCTTAAGTC 759
Db 307 -----AenglyValSer- 310
QY 760 CAANAATCAAAACATTTTGCTACTGATAGTGGCGCATGTCAATAAAGCTTGAGAAAGCTG 819
Db 310 ----- 310
QY 820 ACTTACTAAAGGCTATTCAAGACAAATGATGCTAAAGCTCAAGCTTAACAGCACTACT 879
Db 311 -----CysGlnGluThrAlaValThrlGlnThrTyrglyLeuSerAsnG 326
QY 880 TTGAG-----GTCAATTGATTGCAAGCGATGCAACATTACTGATGCAAAACGGCAAG 933
Db 326 LysGluProCysValLeuProPheThr-----TyraGlnGlyArgT 339
QY 934 TCTACTTTGCT-----GACAAAGATGGTTCGGTAACTTGGCCGACCCCAAC 978
Db 339 hrrPheTySerCysThrThrlGluGlyArgGlnAspGlyHisleuTrpCysSerThrThrs 359
QY 979 CTGTCCAGAAATTTTGTCTAAGCGGCAATGTGCGGTTAGACCATTAAGAAACCA 1038
Db 359 eraAsnTyrglu----- 362
QY 1039 TACAAAAACCAAGCAAAATCTGTGTGATGTGGAATATAGTGAAGTTTACTCCCTTAAC 1098
Db 363 --GlnAspGlnLeuTySerPheCysThrAspHisThrVal----- 375
QY 1099 CTGATGACGATTTTCAAGACCAAGCTCAAAAGATTAAGCTAATTGAAACACTAGCTATG 1158
Db 375 ----- 375
QY 1159 GTGACACATCATCATCTCAAGAAATTACTAGCTCAAGACCAAGCATTTTAAACAAAAC 1218
Db 376 -----LeuValGlnThrGlnGly----- 381
QY 1219 ACCGAGGTATACGATTTATGACACGTGACTCTCATGCTGATCATGACATGACATTT 1278
Db 382 -----GlyAsnSerAenglyAlaleuGlyHis----- 390
QY 1279 TCCGTACGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAG 1338
Db 391 -----PheProPheLeuTyraAsnHis----- 398
QY 1339 CTTATAGGATCATTAATAAATCTGCTGATGAAAGAAATTAACAACAGCTGATCT 1398
Db 399 -----AsnTyThrAspCysThrS 405
QY 1399 CTGAGAAATATTAAGTCTCTTAATAAAGGGAAGAAACCGTATGATCCCTTGAATCAAGTC 1458
Db 405 erGluGly-----ArgAlaAspA 411
QY 1459 ACTTGAAA--CTGTTCACCATCAAAATAGCTGATGTCGATCAAC--G 1503
Db 411 smMetLysTrpCysGlyThrThrGlnAsnTyraAspAlaAspGlnLysPheGlyPheCysP 431
QY 1504 AATTGCTTAAAAAGTGAAGCTCTTAACAAGCTAAGGAAAGTAATTAGACTTCAAGAGATT 1563
Db 431 rometAlaAlaHisGlnGluIleCysThrThrAsnGlu----- 443
QY 1564 TATACGATCTCGTGATAGGCTAAAGCTAACAACAATCTCGATGCTTTTGGTATTA 1623
Db 444 -----GlyValM 446
QY 1624 TGAATATACCTTAATCTGAAAAGTGAAGATTAATCAAGATGACCAACCGTATCATTA 1683
Db 446 et--TyraArgGlnGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgC 465
QY 1684 CCGTTTATATGGGCAAGGCAAGCCGGAAGAGAAATGCTAGCTAATTAAGCTGGTGGTG 1743
```

```
Db 465 yThrCysValGlyAsn-----GlyArgG 473
QY 1744 GCCAGCGGCAACAGATTGA-----CCATAGCTGAAGAAAGTGTGATCAATGCTGCTG 1797
Db 473 LysGluThrTrpCysIleAlaTySerGlnLeuAspGlnCysIle-----ValAspA 491
QY 1798 GGACTTCTATGTGTGCGGAGAAAGTGGAGAGGCGCTTACCAAGCTGATGATGAGTAG 1857
Db 491 spLleThrTyraAsnValAsnAspThrPheHisLysAspHisGlnGluGlyHisMetLeuA 511
QY 1858 ATTGTACTTGCCCTGGGAGAAAGGCAAGCGCATCACTTGCACTTGAATAATGATGCA 1917
Db 511 encYserThrCysPheGlnGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysG 531
QY 1918 ACCATGAGACACACAGACATCTTAATTAATTTGAACACCTGAGACAGAAAGATATC 1977
Db 531 LAspSerGlnThrGlnCysTyrglyArgGlyLeuGlyAspSerTrpGlnLysTyraValHis 550
QY 1978 GAGGAACCTGCTCCAGTGCATCTGACAGAGCAACGCGGAGAGAGTGAAGTGAAGA 2037
Db 551 --GlyValArgTyrglnCysTyrglyArgGlyLeuGlyAspSerTrpGlnLysTyraValHis 569
QY 2038 GGCACACTGTGTGACAGACCATGCGAGCGAATCTGGGCCC 2078
Db 570 -----ProLeuGlnThrTyrglyProSerSerGlyPro 580

RESULT 9
US-09-961-403-1
; Sequence 1, Application US/09961403
; Publication No. US20030077589A1
; GENERAL INFORMATION:
; APPLICANT: HR-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-961-403-1

Alignment Scores:
Pred. No.: 7,49e-49 Length: 2386
Score: 638.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
Gaps: 28

US-09-940-235-12 (1-2096) x US-09-961-403-1 (1-2386)
QY 5 GACCATTCATGTTGTTGCTCAGTGCAGACGTTTGGACAGACAGTGCCTTACGT--- 61
Db 122 AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 141
QY 62 --TCGTCGCGTATCGTGATTAATTTGCTAACAACATTAAGGCAAC 106
Db 142 HisGlnGlyGlnAsnSerTyraLysIleGlyAspThrTrp-----ArgArgPr 157
QY 107 CCGCCAGCTTAGCCGGGTCTTCAACAGACAGACAGATCATGCGACCCGTGGCCAGGA 166
Db 157 ohIsGlnThr----- 160
```


QY 167 CCCAAGCGTCCGAGATCTCGATCCGCGAATTAATGACTCACTATAGGAGACA 226
 Db 160 -----
 QY 227 CAACGGTTCCCTAGAA--ATAATTTGTTTAACTTTAAGAGAGATATCCATGG 283
 Db 161 -GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGlu-TyrThr----- 178
 QY 284 TGCAGACAAAGATTGTACCAATAGCTAGAGAGTTTGTATCATGCTGGTGGGACTT 343
 Db 179 -----CysLysProIleAlaGluLysCysPheAspHisAlaIleGlyThrS 194
 QY 344 CCTATGTGTGGAGAAACGTGGGA----- 368
 Db 194 eTyrValValGlyGluThrTTPGluLysProTyrGlnGlyTyrMetMetValAspCysT 214
 QY 369 -----GAGGACGCGAAGCATCTGACCTTGTAGAAATAGATGCAACGATC 417
 Db 214 hCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnArg 234
 QY 418 AGGACCAAGGAGATCTATAGAAATTTGGAGACCTTGGACGAAGAAGATATGAGAGA 477
 Db 234 LAspThrArgLysThrTyrArgIleGlyAspThrTTPSerLysLysAspAsnArgGlyA 254
 QY 478 ACCTGCTCAGTCATCTGCACAGGCAACGCGAGAGAGATGGAGATGTGAGAGACA 537
 Db 254 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTTPryCysGlnArgHisT 274
 QY 538 CCTGTGTGACAGACCAATCGAGCGGATCTGGCCCTTACCGATGTTGTAATGCT--- 593
 Db 274 hSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaIleValT 294
 QY 593 ----- 593
 Db 294 yGlnProGlnProHisProGlnProProTyrGlnHisCysValThrAspSerGlyV 314
 QY 594 -----GGACCTGAGTGGCTGTAGACGCTCATCTGTCAACAAGCCGCAAT 639
 Db 314 aValTyrSerValGlyMetGlnTTPLeu-----LysThrGlnGlyAsnLysGlnM 331
 QY 640 TGGTGTGATGCGTGTGCTACTGTTGAGGGGAGCAATCAAGACATTGCTTAATTTT 699
 Db 331 eLeuCysThrCysLeuGly----- 337
 QY 700 TTGAAATCGATTCAATCAACGCTGCTCATGGAGAGAAAGACAGCAAGCTTAAGTC 759
 Db 338 -----AsnGlyValSer- 341
 QY 760 CAAATCAAAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTG 819
 Db 341 ----- 341
 QY 820 ACTTAATAAGGCTATTCAAGAAATTTGATCGCTAACGTCACAGTAAGCAGCACTACT 879
 Db 342 -----CysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerArgS 357
 QY 880 TTGAG-----GTCAATTGATTTTGCAGCGATCAACCATTAATGATGAAACGCGAAG 933
 Db 357 LysLysProCysValLeuProPheThr-----TyrAsnGlyArgT 370
 QY 934 TCTACTTTGCT-----GACAAAGATGTTGGTAACTTGGCCGACCAAC 978
 Db 370 hPheTyrSerCysThrThrGlnGlyArgGlnAspGlyHisLeuTTPryCysSerThrThrS 390
 QY 979 CTGTCCAAAGATTTTGTAAAGGAGCATGTGGCGCTTAGACCATTAATAAGAAAAACAA 1038
 Db 390 eAsnTyrGlu----- 393
 QY 1039 TACAAACCAAGCAAGATCTGTGATGTGGAATATATCTGACATTACTCCCTTAACC 1098
 Db 394 --GlnAspGlnLysTyrSerPheCysThrAspHisThrVal----- 406
 QY 1099 CTGATGACGATTTGACAGCAGGCTGTCAAAAGATACTAAGCTATTGAAAAACACTAGCTATCG 1158

Db 406 ----- 406
 QY 1159 GTGACACATGACATCTCAAGAAATTTACTAGCTCAAGACCAAGCATTTTAAACAAAAAC 1218
 Db 407 -----LeuValGlnThrGlnGly----- 412
 QY 1219 ACCAGGCTATACGATTTATGAGCGTACTCTTCAATGCTCACTGATGCAATGACATTT 1278
 Db 413 -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
 QY 1279 TCCGTACGATTTTACCAATGATGACAGATTTTACTTACCTGTTTAAAAATCGGGAACAG 1338
 Db 422 -----PheProPheLeuTyrAsnHis----- 429
 QY 1339 CTATAGGATCAATAAAAATCTGCTGTAATGAAAGAAATTAACAACATGACCTGATCT 1398
 Db 430 -----AsnTyrThrAspCysThrS 436
 QY 1399 CTGAGAAATATTAGCTCTTAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATGCCAGTC 1458
 Db 436 eArgLysGly-----ArgArgAspA 442
 QY 1459 ACTTGAA-----CTGTTCACATCAAAATACGTTGATGTGATACCAAC-----G 1503
 Db 442 snMetLysTTPryCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysP 462
 QY 1504 AATTGCTAAAAAGTGAGACGCTTTAACAGCTAGCGAAAGTAACTTAGACTTCAGAGATT 1563
 Db 462 rMetAlaIleHisGlnGluThrLecCysThrThrAsnGlu----- 474
 QY 1564 TATAGATCTCTGATGATGAGCTTAACCTATCTACAAACATCTGATGCTTTGGTATTA 1623
 Db 475 -----GlyValM 477
 QY 1624 TGACTATACCTTAATGGAAGAAAGTAAATACAGATGACACCAACGATACATTA 1683
 Db 477 eC--TyrArgIleGlyAspGlnTTPAspLysGlnHisAspMetGlyHisMetMetArgC 496
 QY 1684 CCGTTTATGATGGGCAAGGCAACCGAAGAGAAATGCTAGTCACTTAGCTGGTGGTG 1743
 Db 496 yThrCysValGlyAsn-----GlyArgS 504
 QY 1744 GCCAGGCGCAACAGATTGTA-----CCCATGCTGAGAGTGTTTGATGATGCTGCTG 1797
 Db 504 LysLysTTPThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIle-----ValAspA 522
 QY 1798 GGACTTCTATGTGTGGAGAAAGCGGAGAAAGCCCTTACCAAGGCTGATGATGATG 1857
 Db 522 ePileThrTyrAsnValAsnAspThrPheHisLysArgHisGlnGluGlyHisMetLeuA 542
 QY 1858 ATTGTAATGCTGCTGGAGAGGAGCGAGCGGACATCACTTGTGCAATTTGAATAGATGCA 1917
 Db 542 snCysThrCysPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgG 562
 QY 1918 ACATCAAGACACAAAGACATCTATAGATTTGAGACACCTGAGCAAGAAGATTAATC 1977
 Db 562 LAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTTPGluLysTyrValHis- 581
 QY 1978 GAGGAAACCTGCTCAATGATCTGACAGGCAACGCGGAGAGAGATGGAAGTGTAGA 2037
 Db 582 --GlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTTPHisCysGln- 600
 QY 2038 GGACACCTGTGTCAGACACCATGAGCGGATCTGGCCCC 2078
 Db 601 -----ProLeuGlnThrTyrProSerSerSerGlyPro 611
 Db 611 -----

RESULT 10
 US-10-788-792-206
 ; Sequence 206, Application US/10788792
 ; Publication No. US20040191819A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bayer Pharmaceuticals Corporation


```

Db      390  eTAsnTyrglu----- 393
QY      1039  TACAATAACCAAGCGAAATCTGTGATGTGGAAATATATCTAGAGTTACTCCTTAAC 1098
Db      394  --GlnAspGlnIuTySerPheCysThrAspHisThrVal----- 406
QY      1099  CTGATGACGATTTGAGACCAAGGCTCAAGATGACTATGAGTATGAAAACAGTAGTATCG 1158
Db      406  ----- 406
QY      1159  GTGACACATCATCTCAAGAAATTACTAGCTCAAGACCAAGCATTTAAACAAAACC 1218
Db      407  -----LeuValGlnThrGlnGly----- 412
QY      1219  ACCGAGCTATACGATTTATGACGAGTCTCTCATGCTCATGACATGACATTT 1278
Db      413  -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
QY      1279  TCCGTCAGATTTTACCAATGATCAAGAGTTTACTTACCGTGTAAAAATCCGGAAACAG 1338
Db      422  -----PheProPheLeuTyraAsnAsnHis----- 429
QY      1339  CTATATGATCATATAAAATCTGTGATGATGAAGAAATAAACAACACTGACCTGATCT 1398
Db      430  -----AsnTyThrAspCysThrS----- 436
QY      1399  CTGAGAAATATATCTCTTAAAAAAGGGAAGCCGTATGATCCCTTTGATCGCATC 1458
Db      436  erGluGly-----ArgArgAspA 442
QY      1459  ACTGGAAT---CTGTTCAACATCAAAATACGTTGATGTGATACCAAC-----G 1503
Db      442  smteLysTrpCysGlyThrThrGlnAsnTyraAspAlaAspGlnLysPheGlyPheCysP 462
QY      1504  AATGCTAAAAAGTAGAGAGCTCTTAACAGCTAGCGAAAGTAACTTAGACTTCAGAGATT 1563
Db      462  rometAlaAlaHisGluGlnIuIeCysThrThrAsnGlu----- 474
QY      1564  TATACAGCTCTGTGATTAAGCTAAACTACTTCAACAACATCTCGATGCTTTGGTATTA 1623
Db      475  -----GlyValIn 477
QY      1624  TGGACTATACCTTAACCTGAAAAGTAGAGATATACAGATGACACCAACCGTATCATTA 1683
Db      477  et---TyArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetCysC 496
QY      1684  CCGTTATATAGGGCAAGGCAAGCCGAAGAGAGATGTAGTACCATTTAGCTGGTGGTG 1743
Db      496  ystThrCysValGlyAsn-----GlyArgG 504
QY      1744  GCCAAGCGCAACAGATTGTA-----CCCATAGCTGAGAAAGTGTTTGATCATGCTGCTG 1797
Db      504  LysIuTrpThrCysIleAlaTySerGlnLeuArgAspGlnCysIle-----ValAspA 522
QY      1798  GGACTTCTATGTGTGTGGAGAAACGTGGAGAAACCCCTACCAAGGCTGAGATGATG 1857
Db      522  spIeThrTyraAsnValaAsnAspThrPheHisLysArgHisGluGlnGlyHisMetLeu 542
QY      1858  ATTGACTTGGCTGGAGAGAGAGCGAGCGAGCATCATCTTGACCTTCAAAAATGATGTA 1917
Db      542  snCysThrCysPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 562
QY      1918  ACGATCAGACAGACAGACAGATCTTATGAAATTTGAGACACCTGAGACAGAAAGATTAAT 1977
Db      562  LnaAspSerGlnThrGlyThrPheTyrglnIleGlyAspSerThrGlnLysTyrglnHis 581
QY      1978  GAGGAACCTCTCCAGTGCATCTGCACAGCGCAAGCGCCGAGAGAGATGGAAGTGTGAA 2037
Db      582  --GlyValArgTyrglnCysTyrglyArgGlyArgGlyArgGlyArgGlyArgGlyArg 600
QY      2038  GGACACCTCTGTGACAGACAGATGAGGAGATCTGGGCC 2078
Db      601  -----ProLeuGlnThrTyrrProSerSerSerGlyPro 611

```

```

RESULT 12
US-10-485-758-4
; Sequence 4, Application US/10485758
; Publication No. US20050129681A1
; GENERAL INFORMATION:
; APPLICANT: Varner, Judith A.
; TITLE OF INVENTION: Methods for Inhibiting Angiogenesis
; FILE REFERENCE: US00-07325
; CURRENT APPLICATION NUMBER: US/10/485,758
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: 60/310,645
; PRIOR FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-758-4

Alignment Scores:
Pred. No.: 7,49e-49 Length: 2386
Score: 638.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
Gaps: 28

US-09-940-235-12 (1-2096) x US-10-485-758-4 (1-2386)
QY      5  GACCATCATGTTGTTGCTCAGTCCGACAGCGTTTGGACAGACGATCGCTTACGTT--- 61
Db      122  AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 141
QY      62  -----TCGCTCGGTATTCGGTGAATTCATTCGCTAACCAATTAAGCGAAC 106
Db      142  HisGluGlyGlyGlnSerTyrr-LysIleGlyAspThrTrp-----ArgArgP 157
QY      107  CCGCAGCCTTAGCCGGGATCCTCAACGACAGACAGATCATCGCACCCGTGGCCAGA 166
Db      157  ohisGluThr----- 160
QY      167  CCCAAGCTGCCCCAGATCTCGATCCCGGAAATTAATACACTCACTATAGGAGACCA 226
Db      160  ----- 160
QY      227  CAACGGTTCCCTTAGAA---ATAATTTGTTAACTTTAAGAGAGATATACATG 283
Db      161  -GlyGlyTyrrMetLeuGlnCysValCysLeuGlyAsnGlyLysGlyGlu-TripThr--- 178
QY      284  TGCACACCAACAGATTGTACCCATAGCTGAGAAAGTGTTTGATCATGCTGGGCACTT 343
Db      179  -----CysLysProIleAlaGlnLysCysPheAspHisAlaGlyHis 194
QY      344  CCTATGTTGCTGGAGAAACGTGGA----- 368
Db      194  erTyrrValValGlyGluThrTrpGlnLysProTyrglnGlyTrpMetMetValAspCys 214
QY      369  -----GAAGCAGCGGACCGATCATCTTGCACTTTAAGAAATGATGACAGATC 417
Db      214  hrcysLeuGlyGlnGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 234
QY      418  AGGACACAGACATCTTATGAAATTTGAGACACCTGAGACAGAAAGATTAATGAGAA 477
Db      234  LnaAspThrArgThrSerTyraGlyIleGlyAspThrTrpSerLysAspAsnArgGly 254
QY      478  ACCTGTCCAGTGCATCTGCACAGCGCAAGCGCCGAGAGAGATGGAAGTGTGAGGACA 537
Db      254  snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlnLysCysGlyLysGlyHis 274
QY      538  CCTGTGTGACAGACATATGAGGAGATCTGGCCCCCTTACCGAATGTTGATTTGCT--- 593

```

```

Db 214 htservalglnthrhSerSerglySerGlyProPheThrAspValArgAlaValr 294
Oy 593 -----
Db 294 yrglnProglInProhIsProglInProProProTyrglyhIsCySvalThAspSerglyv 314
Oy 594 -----GACCTGAGTGGCTGTGAGACCGTCATCTGTGCACACAGCCAT 639
Db 314 aIvalTySerValGlyMetGlnTrpLeu -----LysThrGlnGlyAsnLysGlnm 331
Oy 640 TGGTTGTTACCGTGGCTGTGACTGTGTGAGGGAGCAATCAAGACATTAGCTTAAATTT 699
Db 331 etLeuCyThrCySleugly ----- 337
Oy 700 TTGAATCGATCTAACATACGACGACCTGCTCATGAGAGAAAGACAGACGCTTAATGC 759
Db 338 -----AsnGlyValSer- 341
Oy 760 CAAAATCAAAACCATTTGCTACTGATAGTGGCGCATGTGCATAAACTTGAGAAAGCTG 819
Db 341 ----- 341
Oy 820 ACTTACTAAAGCTATTCAAGAACAATTGATCGCTAACGTCACAGTAACGACGACTACT 879
Db 342 -----CySglnGlnThrAlaValThrhGlnThrTyrglyGlyAsnSerAsng 357
Oy 880 TTGAG-----GTCATGATTTTGGACAGCATCAACCATTAATGATCGAAACGGCAGG 933
Db 357 lylGlnProCySvalLeuProPheThr-----TyraSnglyAArgT 370
Oy 934 TCTACTTTGCT-----GACAAAGATGGTTCGTAACTTCGACCCGACCAAC 978
Db 370 hrPheTySerCySthrThrhGlnGlyArgGlnAspGlyhIsLeuTrpCySerThrThrs 390
Oy 979 CTGTCCAGAAATTTTGTCTAAGCGACATGTGCGCTTGAACCATTAAGAAACCA 1038
Db 390 ezaSntyrcGlu----- 393
Oy 1039 TACAAACCAAGCGAAATCTGTGATGTGAATATATCTGTACAGTTTACTCCCTTAACC 1098
Db 394 --GlnAspGlnLysTrpSerPheCySthrAspHisThrVal----- 406
Oy 1099 CTGATGACGATTTTCAGACCAAGCTCAAGATATCTAAGCATTTGAAACACTAGCTATCG 1158
Db 406 ----- 406
Oy 1159 GTGACACCATCATCTCAAGATTTACTAGCTCAACGACACAAACATTTTAAACAAAACC 1218
Db 407 -----LeuValGlnThrhGlnGly----- 412
Oy 1219 ACCCAGGCTATACGATTTATGAACGTGACTCTCAATCGTCACTCATGACATGACATTT 1278
Db 413 -----GlyAsnSerAsnGlyAlaLeuCySHis----- 421
Oy 1279 TCCGTAAGATTTTACCAATGATCAAGATTTTACTTACCGTGTAAATAACGGGAACAAG 1338
Db 422 -----PheProPheLeuTyraAsnAsnHis----- 429
Oy 1339 CTTATAGATCAATAAAATCTGTCTGAATGAGAAATAAACAACACTGACCTGATCT 1398
Db 430 ----- 436
Oy 1399 CTGAGAAATATTACGCTTAAAGGGGAAAGCGGTATGATCCCTTGTGATCGCATC 1458
Db 436 erGlnGly-----ArgArgAspA 442
Oy 1459 ACTTGA---CTGTTCACCATCAATACGTTGATGTGATACCAAC-----G 1503
Db 442 smMeLysTrpCySglTyThrhGlnAsnTyraSnpAlaAspGlnLysPheGlyPheCyS 462
Oy 1504 AATTGCTTAAAGAGTGAAGCTCTTAAACAGCTAGGACGTAAGTAAGTGAAGAT 1563
Db 462 rometalAlaHisGlnGlnLysCySthrThrasnGlu----- 474

```

```

Oy 1564 TATACATCTCGTATAGAGCTAAACTACTCTAACACATCTCGATCTTTGGTATTA 1623
Db 475 -----GlyValm 477
Oy 1624 TGAATATACCTTAACCTGAAAGAGTAAGATTAACATGATGACACCAACGCTATCATTA 1683
Db 477 et---TyraGlnGlyAspGlnTrpAspLysGlnHisAspMeGlyhIsMetMetAlaGc 496
Oy 1684 CCGTTATATGGGCAAGCAGACCCGAAAGAGAAATGCTAGTACATTAGCTGTGTGTG 1743
Db 496 yethrCySvalGlyAsn-----GlyArgG 504
Oy 1744 GCCAGGCGCAACATTTGTA-----CCATAGCTGAGAGTGTGTTGATCATGTGCTG 1797
Db 504 lylGlnTrpThrCySleAlaTySerGlnLeuAlaAspGlnCySle-----ValAspA 522
Oy 1798 GGAATCTCTATGTGTGTGAGAAACGTGGAGAAAGCCCTTACCAAGGCTGTGATGATG 1857
Db 522 eplIethrTyraSnpValAsnAspThrPheHisLysValGlnGlnGlyhIsMetLeuA 542
Oy 1858 ATTGTACTTGTGCTGGAGAAAGCGACGACGATCATCTTGCACTTTAGAAATAGATCA 1917
Db 542 snCySthrCySPhelGlyGlnGlyArgGlyArgTrpLysCyAspProValAspGlnCySg 562
Oy 1918 ACATCAGACACACAGACATCTTATGAATTTGAGACACCTGAGACAAAGATATATC 1977
Db 562 lnsPserGlnThrhGlyThrhPheTyrglnGlyAspSerTrpGlnLysTyraValHis 581
Oy 1978 GAGAAACCTGCTCTCACTGATCTGACACAGGCAACGCGGAGAGAGTGAAGTGTAGA 2037
Db 582 --GlyValaLysTyrglnGlyCySthrTyrglyArgGlyhIsGlnTrpHisCySgln- 600
Oy 2038 GGCACACCTGTGTCAGACACCATGACGCGGATTTGGCCCC 2078
Db 601 -----ProLeuGlnThrTyraProSerSerSerglyPro 611

RESULT 13
US-10-485-758-9
; Sequence 9, Application US/10485758
; Publication No. US20050129681A1
; GENERAL INFORMATION:
; APPLICANT: Varmer, Judith A.
; TITLE OR INVENTION: Methods for Inhibiting Angiogenesis
; FILE REFERENCE: USCD-07325
; CURRENT APPLICATION NUMBER: US/10/485,758
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: 60/310,645
; PRIOR FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2386
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-485-758-9

Alignment Scores:
Pred. No.: 7,49e-49 Length: 2386
Score: 636.50 Matches: 208
Percent Similarity: 36.69% Conservative: 69
Best Local Similarity: 27.55% Mismatches: 151
Query Match: 17.14% Indels: 328
DB: 5 Gaps: 28

US-09-940-235-12 (1-2096) x US-10-485-758-9 (1-2386)
Oy 5 GACATTCATGTTGTTGCTGAGTGGACAGCTTTTGCAGACAGAGTGGCTTCACT--- 61
Db 122 AspCySthrCySleGlyAlaGlyArgGlyArgGlyLeSerCySthrIleAlaAsnArgCyS 141
Oy 62 -----TGCTCGCGTATCGGTGATTCATTTCTGCTTAACAGTAAGGCAACC 106

```

Db 142 HisGluGlyGlySerTyr-LysIleGlyAspThrTrp-----ArgArgPr 157
 QY 107 CGCCAGCTTACCGGGTCTCTCAACGACAGACGATCATGGCAGCCGGCCAGCA 166
 Db 157 OHISGLUTHR----- 160
 QY 167 CCCAAGCTGCGCCGAGATCTCGATCCGGAAATTAAACACTACTATAGGGAGACCA 226
 Db 160 ----- 160
 QY 227 CAACGTTTCCTCTAGAA--ATAATTTTGTAACTTTAAGAAGAGATATACATGG 283
 Db 161 -GlyGlyTyrMetLeuGlyCysValCysLeuGlyAsnGlyGlyGlyu-TrpThr- 178
 QY 284 TGCAGACAAACAGATTGATACCATAGTAGAAGTGTGTGATCATGCTGGGAGCT 343
 Db 179 -----CysIysProIleAlaGluCysPheAspHisAlaIleGlyThrS 194
 QY 344 CCTATGTGTGTCGAGAAACGTGGGA----- 368
 Db 194 erTyrValValIleGlyuThrTrpGluLysProTyrGlnGlyTyrMetMetValAspCysT 214
 QY 369 -----GAAGCAGCGGACCGCATCATCTTGACCTTGAATAATAGTCAACGATC 417
 Db 214 hrCysLeuGlyGlyGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAspG 234
 QY 418 AGGACAAACAGCATCTTATAGATTTGAGACACCTGGAGCAAGAGATATACGAGAA 477
 Db 234 InAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysAspAsnArgIYA 254
 QY 478 ACCCTCTCAGTGCATCTGACAGAGCAGCAGCAGAGAGAGTGAAGTGAAGGACCA 537
 Db 254 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlyuTrpLysGlyIleArgHisT 274
 QY 538 CCTGTGTGACAGCAACATGACGCGAGTGGCCCTTACCGATGTTCTGATTTGCT- 593
 Db 274 hrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaIleValT 294
 QY 593 ----- 593
 Db 294 yrGlnProGlnProHisProGlnProProTyrGlyHisCysValThrAspSerGlyV 314
 QY 594 -----GACCTGAGTGGTCTGTAGACCGTCCATCTGTCAACACGACCAAT 639
 Db 314 alValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGlnM 331
 QY 640 TGGTTGTAGCGCTGTGCTGCTGTTGAGGGAGCAATCAACATTAAGTCTTAAATTT 699
 Db 331 etLeuCysThrCysLeuGly----- 337
 QY 700 TTGAATGCATCTAACATCAACGACTGCTCATGAGAGAGAAAGACAGACAGGCTTAAGTC 759
 Db 338 -----AsnGlyValSer- 341
 QY 760 CAAATCAAAACCATTTGTCTACTGATGATGGCGCGCATGTCATATAACTTGAGAAAGCTG 819
 Db 341 ----- 341
 QY 820 ACTTACTAAGGCTATTCAAGAACATTTGATGCTTAACGTCAACAGTAAACGACACTACT 879
 Db 342 -----CysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnG 357
 QY 880 TTGAG-----GTCAATGATTTTTCAGAGGAGTCAACCATTTACTGATCGAAACGGCAGG 933
 Db 357 LysGlnProCysValLeuProPheThr-----TyrAsnGlyArgT 370
 QY 934 TCTACTTTGCT-----GACAAAGATGGTTGCGTAACTTGGCCGACACCAAC 978
 Db 370 hrPheTyrSerCysThrThrGluGlyArgGlnAspIleHisLysLeuTyrCysSerThrS 390
 QY 979 CTGTCCAGAAATTTTGTCTAGCAGACATGTGCGCTTAGACCATATTAAGAAACCA 1038
 Db 390 erAsnTyrGlu----- 393

QY 1039 TACAAACCAAGCGAAATCTGTGATGGAATATATCTGACATTTACTCCCTTAAAC 1098
 Db 394 --GlnAspGlnLysTyrSerPheCysThrAspHisThrVal----- 406
 QY 1099 CTGATGACGATTTTACACCAAGTCTCAAAAGATCTAAGCTATTTGAAAACACTAGCTATCG 1158
 Db 406 ----- 406
 QY 1159 GTGACACCATTCATCTCAAGAAATTACTAGTCAAGACAAAGCATTTTAAACAAAAC 1218
 Db 407 -----LeuValGlnThrGlnGly----- 412
 QY 1219 ACCGAGCTATAGATTTTATGACGTGACTCTCAATGCTGACCTGATGACATGATTT 1278
 Db 413 -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
 QY 1279 TCCGTACGATTTTACCAATGGATCAAGGTTTACTTACCGTGTATTAATAATCGGAAACAG 1338
 Db 422 -----PheProPheLeuTyrAsnAsnHis----- 429
 QY 1339 CTATATAGATCATATAAAATCTGGTCTGAATGAAGAAATTAACAACACTGACTGATCT 1398
 Db 430 -----AsnTyrThrAspCysThrS 436
 QY 1399 CTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAAGCGTATGATCCCTTGTATGCGACATC 1458
 Db 436 erGluGly-----ArgArgAsp 442
 QY 1459 ACTTGAAA--CTGTGACCATCAATCAATACGTTGATGTCATACCAAC-----G 1503
 Db 442 snMetLysTrpCysGlyIleThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysP 462
 QY 1504 AATTGCTTAAAGTAGACGAGCTTTAAACAGCTTACCGAAGCTTAATTTAGACTTACAGAGATT 1563
 Db 462 roMetAlaAlaHisGluGluIleCysThrThrAsnGlu----- 474
 QY 1564 TATATGATCCTCGTGAATAGGCTTAATCTATCTACAAACATCTGATGCTTTGGTATTA 1623
 Db 475 -----GlyValM 477
 QY 1624 TGCATATACCTTAACTGAAAAGTAGAGATTAATCAAGATGACACCAACCGTATCANAA 1683
 Db 477 et---TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetClyHisMetMetArgC 496
 QY 1684 CCGTTATATAGGCGAACGACCCGAGAGAGAAATGCTACCTTACATTTAGCTGGGTG 1743
 Db 496 ysThrCysValGlyAsn-----GlyArgG 504
 QY 1744 GCCAGGCGCAACAGATTGTA-----CCCATAGCTGAGAAAGCTTTTGTATCATGCTGCTG 1797
 Db 504 LysGluTrpThrCysIleAlaIleTyrSerGlnLeuAspArgLncysIle-----ValAsp 522
 QY 1798 GAACTTCTATGTGTGCGAGAAAGTGGAGAAAGCCCTTACCAAGGCTGATGATGTAG 1857
 Db 522 splIeThrTyrAsnValAsnAspThrPheHisLysValGlnHisGlnGlyHisMetLeuA 542
 QY 542 snCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysG 562
 Db 1918 ACATCAAGACACAAAGACATCTTATGAATTTGAGAGACACTGAGACAAAGATATC 1977
 QY 562 InAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHis- 581
 Db 1978 GAGGAAACCTGCTCAGATGATCTGACAGGCAACGCGGAGAGAGAGTGAAGTGTGAGA 2037
 Db 582 --GlyValAlaGlyTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGln- 600
 QY 2038 GGCACACCTTGTGACAGACCAATGACGAGCTGGCCCC 2078
 Db 601 -----ProLeuGlnThrTyrProSerSerSerGlyPro 611

```

RESULT 14
US-10-236-392-2
Sequence 2, Application US/10236392
Publication No. US20040067490A1
GENERAL INFORMATION:
APPLICANT: Anderson, David M
APPLICANT: Boldog, Ferenc L
APPLICANT: Burgess, Catherine, E
APPLICANT: Caeman, Stacie J
APPLICANT: Caterton, Elina
APPLICANT: Chapoval, Andrei
APPLICANT: Crabtree, Julie
APPLICANT: Edinger, Shlomit, R
APPLICANT: Ellerman, Karen
APPLICANT: Gerlach, Valerie
APPLICANT: Gorman, Linda
APPLICANT: Grose, William M
APPLICANT: Gusev, Vladimir
APPLICANT: Kekuda, Ramesh
APPLICANT: Larocheille, William J
APPLICANT: Li, Li
APPLICANT: MacDougall, John R
APPLICANT: Malysankar, Uriel M
APPLICANT: Miller, Charles E
APPLICANT: Millet, Isabelle
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patnrajan, Weera
APPLICANT: Pena, Carol A
APPLICANT: Peyman, John A
APPLICANT: Rastelli, Luca
APPLICANT: Reiger, Daniel K
APPLICANT: Rotenberg, Mark E
APPLICANT: Shenoy, Suresh
APPLICANT: Shimkets, Richard A
APPLICANT: Smithson, Glenda
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-442A
CURRENT APPLICATION NUMBER: US/10/236.392
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: US09/540,763
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US60/390,155
PRIOR FILING DATE: 2002-06-19
PRIOR APPLICATION NUMBER: US09/635,949
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US60/318,765
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: US60/357,303
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US60/367,753
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER: US60/369,479
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: US09/659,634
PRIOR FILING DATE: 2000-09-12
PRIOR APPLICATION NUMBER: US60/318,120
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US60/318,130
PRIOR FILING DATE: 2001-09-07
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 794
SOFTWARE: Custom
SEQ ID NO 2
LENGTH: 2320
TYPE: PRT
ORGANISM: Homo sapiens
US-10-236-392-2

Alignment Scores:
Pred. No.: 1,266-48 Length: 2320
Score: 636.00 Matches: 205
Percent Similarity: 36.56% Conservative: 71
Best Local Similarity: 27.15% Mismatches: 154

```

Query Match:	17.07%	Indels:	326
DB:	4	Gaps:	23
US-09-940-235-12 (1-2096) x US-10-236-392-2 (1-2320)			
QY	5 GACCATTCATGTTGTTGCTCAGGTGCGACAGCGTTTGACGACAGACAGTCGCTTACAGT---	61	
DB	87 AapCysrthCysaileglValaeglValaagGlyaaGllseerCysrthIlealaaanaGys	106	
QY	62 -----TCGCTCGGATTCGGTGAATTCATTCCTGTAACAGTAAGCAACC	106	
DB	107 HisgluGlygluGlnserTyr-LysIleglYasprThrP-----ArgaHgr 122	122	
QY	107 CCGCAGAGCTTACCGGGTCTTCAAGCAGAGACAGATCATCGACCCGTGGCAGGA	166	
DB	122 OhIsegluThr-----	125	
QY	167 CCGAAGCGTCCGAGATCTCGATCCCGGAAATTAATACACTCATTAAGGAGACCA	226	
DB	125 -----	125	
QY	227 CAACGGTTCCTCTAGAA--ATATTTGTTTAATTAGAGAGAGATATACATGG	283	
DB	126 -glYglYrYrmetleuGluCysValCysleuGlyaaGlygluGlu-TripThr-----	143	
QY	284 TGCAGACCAACAGATGTGATCCCATATGCTGAGAGAGTGTTCATCTCTGGAACTT	343	
DB	144 -----CysIysProIlealagluCysCysPheAsprHislaIaIaglyThrs	159	
QY	344 CCTATGTGTGGAGAAACGTGGA-----	368	
DB	159 erYrValaValaIgluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCysT	179	
QY	369 -----GAAGGACGCGAGCGCATCTTGACCTTGAAGAAATAGTCAACGATC	417	
DB	179 hrcYsleuGlygluGlnserGlyaaGllleThrcYsrthseraGaaanaGysAsnaAsrg	199	
QY	418 AGGACACAAAGACATCTTATGAATTTGAGACACCTGAGCAAGAAAGATTAATGAGAA	477	
DB	199 InaPrThArghrserTyrarGllleGlyAsprThrTrpserIysIysaPanaaHrglyA	219	
QY	478 AACTGCTCCAGTGCATCTTGCAACGGAACGCCGAGAGAGAGTGAAGTGAAGGCACA	537	
DB	219 snleuLeuGlnCysaileCysrthrGlyaaGlyaaGlygluTrpLysCysgluArHnIst	239	
QY	538 CCTGTGTCAAGACCATGAGCGGATCTGGCCCTTCACCGATGTTGATTTGCT----	593	
DB	239 hrserValGlnThrTrpsererGlyserGlyProPheThrAspValaIaIaValT	259	
QY	593 -----	593	
DB	259 yrGlnProGlnProHisProGlnProProTyrGlyHisCysValaThrAspserGlyY	279	
QY	594 -----GACCTGAGTGGCTGCTAGACCGCTCCATCTGTCAACAACAGCCAAAT	639	
DB	279 alValYrserValaIgluMetGlnTrpLeu-----LysThrGlnGlyaaLeuYsGlnM	296	
QY	640 TGGTTGTAGCGCTTCTGCTGACTGTGTAAGGGGACGAATCAAGCATTAAGCTTAAATTT	699	
DB	296 etLeuCyserThrcYsleuGlnaGlyaaGlyaaLysCysGlnGluThrIaIaValaThrGlnThrT	316	
QY	700 TTGAATGATCTTAATCATCAGACCT-----GCTCAAGAGAGAAAGACAG	744	
DB	316 yrGlyGlyaaLeuLeuanaGlnGlyuProCysValaLeuProheThrYraaGlnGlyaaGThPr	336	
QY	745 AGCAAGGCTTAAGTCCAAAATCAAAACCAATTTGCTATGATAGTGCGCGGATGTACATA	804	
DB	336 heYrYserCyserThrThrGlnGlyaaGln-----	345	
QY	805 AACTTGAAGAAAGCTGACTTACTTAAGGCGTATTCAGAAACAATTTGATCGCTAACGTCACA	864	
DB	346 -----AspGlyHisleuThrCysserThrThrs	355	


```

QY      865 GTAACGACGACTTGTGAGTCAATGATTTTGCAGCGATGCAACCATTAATGATGCA  924
      |||||  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      355 eAaSnTyrgluInlYarSglnlyrSerPheCystrhrAspHisThrValGln  374
QY      925 ACGGCAAGGTCTTACTTTGCTGCAAAAGATGGTTCGGTAACCTTGGCCGAACTGGCC  984
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      375 Thr-GlnGly-----GlyAaSerAsnGlyValLeuCyS--  385
QY      985 AAGATTTTGTGTAAGCGGACATGTCGGCTTAGACCATTAAGAAAAACCAATACAA  1044
      385 -----
QY      1045 ACCAAGCGAAATCTGTGATGATGATATATCTGTACAGTTTACTCCCTTAACCTGATG  1104
      385 -----
Db      385 -----
QY      1105 ACGATTTGACACGAGTCTCAAGATATTAAGCTATTGAAACAATAGTATGGTGACA  1164
      385 -----
Db      385 -----
QY      1165 CCATCATCTCTCAAGAAATTAAGTCTCAAGCACAAGCAATTTTAAACAAACCCACG  1224
      385 -----
Db      385 -----
QY      1225 GCTATACGATTTATGAAAGTGACTCTCAATGCTGACTATGACATTAATTTCCGTA  1284
      386 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      386 -----
QY      1285 CGATTTTACCAATGATCAAGAGTTTACTACCGTGTAAAAATCGGGAACAAGCTTATA  1344
      389 eLeuTyrrAsnAsn-----His-AsnTyrrThrAspCystrhrSerGlu  393
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      393 -----
QY      1345 GGATCAATTAATAAAATCTGCTGATGAAAGAAATTAACAACACTGACTGATCTTGAGA  1404
      394 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      394 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1405 AATATTAGCTCTTAATAAAAGGGAAGAACCCGATATGCTTGTATGCGAGTCACTTA  1464
      403 1y-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      403 1y-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1465 AA---CTGTTTCAACATCAATACGTTGATGTCGATCAAC-----GAATTCG  1509
      409 yStrpCySglnYthrGlnAsnTyrrAspAlaAspGlnlyrPheCySProMeTa  429
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      429 yStrpCySglnYthrGlnAsnTyrrAspAlaAspGlnlyrPheCySProMeTa  429
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1510 TAAAAAGTGACGACTTTTAACAGCTAGCGAACTTAATTAAGCTTCAAGATTATAGC  1569
      429 1aAlaHisGlnGluIuLeCystrhrThrAsnGlu-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      429 1aAlaHisGlnGluIuLeCystrhrThrAsnGlu-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1570 ATCCCGATGATAGGCTAAACTACTTACAAACATTCGATGCTTTGGTATTATGACT  1629
      440 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      440 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1630 ATACCTTAACGAAAAAGTAGAGATATACAGATGACACCAACCGTATATGACGTTT  1689
      443 yrrArgIleGlyAspGlnYrrAspGlnlyrAspMetGlnHisMetMetAspCyThrC  463
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      463 yrrArgIleGlyAspGlnYrrAspGlnlyrAspMetGlnHisMetMetAspCyThrC  463
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1690 ATATGCGCAGCGACCCGGAAGAGAGATGCTAGTCAATTAAGCTGTGTGGCCAGG  1749
      463 ySvalGlyAsn-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      463 ySvalGlyAsn-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1750 CGCAACAGATTGTA-----CCCATAGCTGAGAAAGTTTGTATCATGCTGCGGAGACT  1803
      471 rPrhrCysIleAlaTyrrSerGlnLeuArgAspGlnCysIle-----ValAspAspIle  489
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      471 rPrhrCysIleAlaTyrrSerGlnLeuArgAspGlnCysIle-----ValAspAspIle  489
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1804 CTTATGTGTCGGAAGAAACGTCGGAAGAACCCCTAACAGCTGATGATGATGATTTGTA  1863
      489 hrTyrrAsnValAsnAspThrhrPheHisGlyArgGlnHisMetMetLeuAspCys  509
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      489 hrTyrrAsnValAsnAspThrhrPheHisGlyArgGlnHisMetMetLeuAspCys  509
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1864 CTTCCTGAGAGAGAGCGAGCGGACGATCACTTCTAGAAATAGATGCAAGATC  1923
      509 hrCyrrPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg  529
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      509 hrCyrrPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg  529
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1924 AGGAACACAAAGACATCTTATAGAAATTTGGAGACACTGGAGCAAGAAAGTAATCGAGAA  1983

```

```

Db      529 eArgIuThrGlnYthrPheTyrrGlnIleGlyAspSerTrpGluTyrrValHis---GlyV  548
      ::|||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1984 ACCTGCTCCAGTCATCTGCAACGCAACGCGGAGAGAGTGGATGAGAGGACCA  2043
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      548 alArgTyrrGlnCyStryCyStryrGlyArgGlyIleGlyGluTrpHisCySgln-----  565
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      2044 CCTGTGCGACGACATCGAGCGGAGATCTGGCCCC  2078
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      566 --ProLeuGlnIuTryrProSerSerSerGlyPro  576
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 15
US-10-171-311-64
; Sequence 64, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamackar, Shubhangi
; APPLICANT: Glatc, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 2328
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-171-311-64

Alignment Scores:
Pred. No.: 1,26e-48 Length: 2328
Score: 636.00 Matches: 205
Percent Similarity: 36.56% Conservative: 71
Best Local Similarity: 27.15% Mismatches: 154
Query Match: 17.07% Indels: 326
DB: 4 Gaps: 23

US-09-940-235-12 (1-2096) x US-10-171-311-64 (1-2328)
QY      5 GACCATTCATGTTGTTGCTCTACAGTTCGACAGCGTTTGCACAGAGTGGCTTCACT  61
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      95 AspCythrCysIleGlyAlaGlyArgGlyArgIleSerCystrhrIleAlaAsnArgCys  114
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      62 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      115 HisGluGlyGlyGlnIuTrp-lysIleGlyAspThrTrp-----ArgArgPr  130
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      107 CCGCAGCCTTACGCGGGTCTTCAACGACAGACAGCATGATGCGACCCGTGGCCAGA  166
      130 oHisGluTrp-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      130 oHisGluTrp-----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      167 CCCAAGCCTGCCGAGATCTCGATCCGCGCAATTAATACGATCTATATGAGAGACA  226
      133 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      133 -----
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      227 CAACGGTTCCCTTGAAG---ATAATTTGTTTAATTAAGAAGAGATATACAGAG  283
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      134 -GlyGlyTyrrMetLeuGluCySvalCyrrLeuGlyAsnGlyIuTrpThr-----  151
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 02:09:11 ; Search time 6.06893 Seconds
(without alignments)
7479.985 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 3726
Sequence: 1 cgaagaccatcatgtgtgtt.....cctcacagatgtcgttag 2096

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 151242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPFO_spool_P/US09940235/runat_27012006_144220_27676/app_query.fasta_1.7708
-DB=Published Applications_AA_New -QFMT=faetan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAJLEN=200000000 -USER=US09940235_QCGN_1_1 @runat_27012006_144220_27676
-LONGLO -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_New.*

- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1984	53.2	414	US-10-631-558-2
2	1928	51.7	413	US-11-032-951-12
3	634.5	17.0	657	US-10-995-561-622
4	634.5	17.0	984	US-10-995-561-629
5	634.5	17.0	2296	US-10-995-561-633
6	634.5	17.0	2355	US-10-995-561-623
7	634.5	17.0	2384	US-10-995-561-627
8	634.5	17.0	2384	US-10-821-234-1545
9	634.5	17.0	2386	US-10-995-561-626
10	632	17.0	259	US-10-631-558-4

11	632	17.0	300	7	US-11-006-119-31	Sequence 31, Appl
12	632	17.0	642	6	US-10-995-561-631	Sequence 631, App
13	278.5	7.5	1341	6	US-10-995-561-621	Sequence 621, App
14	275.5	7.4	693	6	US-10-995-561-632	Sequence 632, App
15	275.5	7.4	1315	6	US-10-995-561-630	Sequence 630, App
16	275.5	7.4	1348	6	US-10-995-561-624	Sequence 624, App
17	260	7.0	1259	6	US-10-995-561-625	Sequence 625, App
18	221	5.9	1286	6	US-10-995-561-628	Sequence 628, App
19	202.5	5.4	847	6	US-10-995-561-634	Sequence 634, App
20	132.5	3.6	1992	7	US-11-013-759-3	Sequence 3, Appl
21	132.5	3.6	1992	7	US-11-013-759-13	Sequence 13, Appl
22	132.5	3.6	2047	7	US-11-013-759-4	Sequence 4, Appl
23	132.5	3.6	2047	7	US-11-013-759-7	Sequence 7, Appl
24	125	3.4	1417	7	US-11-052-554A-8	Sequence 8, Appl
25	119	3.2	2340	7	US-11-052-554A-171	Sequence 171, App
26	116.5	3.1	470	6	US-10-485-517-319	Sequence 319, App
27	116.5	3.1	1647	7	US-11-052-554A-260	Sequence 260, App
28	115.5	3.1	2890	7	US-11-115-639-31	Sequence 31, Appl
29	115.5	3.1	2890	7	US-11-115-639-32	Sequence 32, Appl
30	115.5	3.1	2890	7	US-11-115-639-33	Sequence 33, Appl
31	115	3.1	21	7	US-11-032-951-7	Sequence 7, Appl
32	115	3.1	463	6	US-10-793-626-960	Sequence 960, App
33	113	3.0	801	7	US-11-174-150-29	Sequence 29, Appl
34	113	3.0	801	7	US-11-124-368A-292	Sequence 292, App
35	113	3.0	1765	6	US-10-055-877-140	Sequence 140, App
36	112.5	3.0	1940	6	US-10-055-877-141	Sequence 141, App
37	112	3.0	1448	6	US-10-485-517-212	Sequence 51, Appl
38	110.5	3.0	2314	7	US-11-013-759-11	Sequence 11, Appl
39	108.5	2.9	4913	6	US-10-453-372-1142	Sequence 1142, App
40	108	2.9	4961	6	US-10-453-372-1132	Sequence 1132, App
41	108	2.9	948	6	US-10-485-517-131	Sequence 131, App
42	107.5	2.9	1416	7	US-11-128-059-60	Sequence 60, Appl
43	107	2.9	1494	7	US-11-128-059-78	Sequence 78, Appl
44	107	2.9	2086	7	US-11-128-059-82	Sequence 82, Appl
45	107	2.9				

ALIGNMENTS

RESULT 1
US-10-631-558-2
; Sequence 2, Appl; Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamnara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-631-558-2
Alignment Scores:

Pred. No.: 1,52e-141 Length: 414
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-12 (1-2096) x US-10-631-558-2 (1-414)
QY 588 ATTGCTGACCTGATGGTGTCTAGACCGTCCATCTGTCAACAACAGCCAAATTGTTGTT 647
Db 1 IleaIaGIyProGIuItrPLeuLeuAspArGProSeRValAsnAsnSerGIuLeuValVal 20
QY 648 AGCGTTCGCTGACTGTGGAGGGGAGCAATTCAGACATTAGTCTTAATTTTGAATC 707
Db 21 ServAlaIaGIyThrValGIuGIyThrAsnGIuAspIleSerLeuysPhePheGIuIle 40
QY 708 GATCTAACATCAAGCACTGCTCATGAGGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db 41 AspLeuThrSerArGProAlaHisGIyGIyLysThrGIuGIuGIyLeuSerProLysSer 60
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTA 827
Db 61 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGIuLysAlaAspLeuLeu 80
QY 828 AAGGCTATTCAAGAACATTTGCTTAACGTCCACAGTAAGACGACATCTTGAGGTC 887
Db 81 LysAlaIleGIuGIuGIuGIuLeuLeuLeuAlaSerAsnAspAspIyRheGIuVal 100
QY 888 ATTGATTTTGGACGAGTACCAATTAATCTGTAACGAGGCTACTTGTGTCAC 947
Db 101 IleAspPheAlaSerAspAlaThrIleThrAspAspAsnGIyLysValIyRheAlaAsp 120
QY 948 AAAGATGCTTCGTAACCTTGCAGACCAACCTGTCCAGAAATTTTGTAAAGCGGACAT 1007
Db 121 LysAspGIySerValThrLeuProThrGIuProValGIuGIuPheLeuLeuSerGIyHis 140
QY 1008 GTGGCGGCTTAACCATTAATAAAGAAACCAATACAAACAGCGAAATCTGTATGTCG 1067
Db 141 ValArgValArgProGIyLysGIuLysProIleGIuAsnGIuAlaLysSerValAspVal 160
QY 1068 GAATTAATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGGTCGAA 1127
Db 161 GIuIyTrThrValGIuPheThrProLeuAsnProAspAspPheAspProGIyLeuLys 180
QY 1128 GATTAAGCTATTGAAACCACTAGCTATCGGTGACACCATGACATCTCAAGATTA 1187
Db 181 AspThrLysLeuLeuLysThrLeuAlaIleGIyAspThrIleThrSerGIuGIuLeuLeu 200
QY 1188 GCTCAAGACAAAGATTTTAAACAAACCAACCGGCTATACGATTTATGAACGTGAC 1247
Db 201 AlaGIuAlaGIuAsnIleLeuAsnLysAsnHisProGIyLysThrIleLysGIuArgAsp 220
QY 1248 TCCTCAATCGTCACTGATGACAAATGATTTTCCGATAGATTTTACATGGATCAAGAG 1307
Db 221 SerSerIleValIleHisAspAsnAspIlePheAspThrIleLeuProAspGIuGIu 240
QY 1308 TTTTACTACCGTGTAAAAATCGGAGACAGCTTATAGGATCAATAAAAATCTGCTCG 1367
Db 241 PheThrIyArgValLysAsnArgGIuGIuAlaIyArgIleAsnLysLysSerGIyLeu 260
QY 1368 AATGAAGAAATTAACACATGACCTGATCTTGAGAAATATTAGCTCTTAATAAAGG 1427
Db 261 AsnGIuGIuLysAsnAsnThrAspLeuLysSerGIuLysLysLysValLeuLysLysGIy 280
QY 1428 GAAAGACCGTATGATCCCTTGTGATGACGATCACTTGAACCTTGACCAATCGTT 1487
Db 281 GIuLysProIyArgProPheAspAspAspSerHisLysLysLysLysLysLysLys 300
QY 1488 GATGTCGATACCAAGAAATGCTTAATAAAGTGAAGCTTAAACGTAACGCAACGTAAC 1547
Db 301 AspValAspThrAsnGIuLeuLeuLysSerGIuGIuLeuLeuThrAlaSerGIuArgAsn 320

QY 1548 TTAGACTTCAGAGATTTATAGATCTCGTATAGGCTAAAGCTTAACATCTC 1607
Db 321 LeuAspPheArgAspLeuIyRAspProArgAspLysAlaLysLysLeuIyRAsnAsnLeu 340
QY 1608 GATGCTTTTGGATTTATGACATTAACCTTAACCTGAAAGATAGAGATTAACGATGAC 1667
Db 341 AspAlaPheGIyLysLeuAspIyRThrLeuThrGIyLysValGIuAspAsnHisAspAsp 360
QY 1668 ACCAAGCGTATCAATACCGTTTATATGGGCAAGCAGACCCGAAAGAGAAATGCTAGTAC 1727
Db 361 ThrAsnArgIleIleThrValIyRMetGIyLysArgProGIuGIuAsnAlaSerLys 380
QY 1728 CATTAAGCT 1736
Db 381 HisLeuAla 383
RESULT 2
US-11-032-951-12
; Sequence 12, Application US/11032951
; Publication No US20050282741A1
; GENERAL INFORMATION:
; APPLICANT: Kyristal, Gerald
; APPLICANT: Rabkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003005
; CURRENT APPLICATION NUMBER: US/11/032,951
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: US 09/919,703
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatsSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-11-032-951-12
Alignment Scores:
Pred. No.: 2,38e-137 Length: 413
Score: 1928.00 Matches: 377
Percent Similarity: 98.95% Conservative: 1
Best Local Similarity: 98.69% Mismatches: 2
Query Match: 51.74% Indels: 2
DB: 7 Gaps: 2
US-09-940-235-12 (1-2096) x US-11-032-951-12 (1-413)
QY 588 ATTGCTGACCTGATGGTGTCTAGACCGTCCATCTGTCAACAACAGCCAAATTGTTGTT 647
Db 1 IleaIaGIyProGIuItrPLeuLeuAspArGProSeRValAsnAsnSerGIuLeuValVal 20
QY 648 AGCGTTCGCTGACTGTGGAGGGGAGCAATTCAGACATTAGTCTTAATTTTGAATC 707
Db 21 ServAlaIaGIyThrVal--GIyThrAsnGIuAspIleSerLeuysPhePheGIuIle 39
QY 708 GATCTAACATCAAGCACTGCTCATGAGGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db 40 AspLeuThrSerArGProAlaHisGIyGIyLysThrGIuGIuGIyLeuSerProLysSer 59
QY 768 AAACCATTTGCTACTGATAGTGGCGGATGTCAATAACTTGAGAAAGCTGACTTA 827
Db 60 LysProPheAlaThrAspSerGIyAlaMetSerHisLysLeuGIuLysAlaAspLeuLeu 79
QY 828 AAGGCTATTCAAGAACATTTGATGCTTAACGTCACAGTAAGCACTTGAAGGTC 887

```

Db      |||
80  LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspArgTyrPheGlnVal 99
Qy      888  ATGTATTTTGGCAAGCATGCAACCATATTGATCGAAACGGCAAGGCTCTTGTCTGAC 947
Db      100  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrPheAlaAsp 119
Qy      948  AAAAGATGTTCCGTTAACTTGGCCGACCCAACTGCTCCAGAAATTTTGTCTAAGCGGACAT 1007
Db      120  LysAspGlySerValThrLeuProThrGlnProValGlnGlnIlePheLeuLeuSerGlyHis 139
Qy      1008  GTGCGCGCTTAGACATTAATAAGAAAAACAATAACAACAACGGAATCTGTGATGTG 1067
Db      140  ValAlaGValAlaArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 159
Qy      1068  GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACAGATTTCAGACAGGCTCAAA 1127
Db      160  GluTyrThrValGlnThrPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 179
Qy      1128  GATTAAGCTATTGAAAAACAATAGTATGAGTGAACCATCATCATCTCAAGAAATTACTA 1187
Db      180  LeuThrLysLeuLeuLysThrIleuAlaIleGlyAspThrIleThrSerGlnGlnLeu 199
Qy      1188  GGTCAAGCAACAAGCATTTTAAACAACAACACCCAGGCTATACGATTATGAAAGTAC 1247
Db      200  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGlnArgAsp 219
Qy      1248  TCCCTCAATGCTCATCATGACAAATGACATTTTCCGTAAGCATTTTAAACAAGTCAAGAG 1307
Db      220  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 239
Qy      1308  TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAAATCGTGTG 1367
Db      240  PheThrTyrArgValLysAsnArgGlnGlnAlaTyrArgIleAsnLysLysSerGlyLeu 259
Qy      1368  AATGAAGAAATAAACAACATGACCTGATCTTGAGAAATAT--TACGTCCTTAAAAA 1424
Db      260  AsnGlnGlnIleAsnAsnThrAspLeuIleSerLeuGlnTyrLysTyrValLeuLysLys 279
Qy      1425  GGGGAAAGCCGATATGATCCCTTGAATGCGAGTCACTGAAATGTTTCAACATCAATAC 1484
Db      280  GlyGlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr 299
Qy      1485  GTTGAATGATATCAACAAGATTGCTAAAAAAGTGAAGAGCTTTAAGCTAGCGAAGCT 1544
Db      300  ValAspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArg 319
Qy      1545  AACTTAGACTTCAAGATTTATACGATCCTCGTATAGGCTTAACTACTCTCAACAAT 1604
Db      320  AsnLeuAspPheArgAspLeuTyrAspProArgAspValAlaLysLeuLeuTyrAsnAsn 339
Qy      1605  CTCGATGCTTTTGGTATTATGACATCTTACTTACGAAAGATAGAGATTAATACGAT 1664
Db      340  LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAsp 359
Qy      1665  GACACAACCGGTATCAACCGTTTATATGAGGCAAGCGACCGCAAGAGAAGATTCCTAC 1724
Db      360  AspThrAsnArgIleIleThrValLysMetGlyLysArgProGlnGlnLeuAlaSer 379
Qy      1725  TACCAT 1730
Db      380  TyrHis 381

```

```

; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-622

Alignment Scores:
Pred. No.: 1.94e-40      Length: 657
Score: 634.50           Matches: 207
Percent Similarity: 36.69%      Conservative: 70
Best Local Similarity: 27.42%    Mismatches: 151
Query Match: 17.03%           Indels: 328
DB: 6                      Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-622 (1-657)
Qy      5  GACCATCATGTTGTGTGTCAGGTGGAGAGCTTTTGGACAGAGTGGCTTCAAGT--- 61
Db      122  AspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 141
Qy      62  -----TGCCTGCGCTATCGGTGATTCATTTCTGTAAACAGTAAAGCAAC 106
Db      142  HisGlnGlyGlnGlnSerTyr-LysIleGlyAspThrTrp-----ArgArgPr 157
Qy      107  CCGCCAGCCTTACCGGGGTCTCAACAGACAGACATCATGCGCACCCGTGGCCAGGA 166
Db      157  ohIeGlnThr----- 160
Qy      167  CCCAAGCCTGCCAGATCTCGATCCCGCGAAATTAATAGATCTATAGGAGACCA 226
Db      160  ----- 160
Qy      227  CAACGGTTCCCTAGAA---ATAATTTGTTTAACTTTAAGAGAGATATACATCG 283
Db      161  -GlyGlyTyrMetLeuGlnCysValCysLeuGlyAsnGlyLysGlnGln-TrpThr--- 178
Qy      284  TGCAAGCAACAAGATTGTATCCCATAGCTGAGAAAGTTTTGTATCATGCTGCTGGAATT 343
Db      179  -----CysLysProIleAlaGlnLysCysPheAspHisAlaIleGlyHis 194
Qy      344  CTTATGTGTCGAGAAACGTGGA----- 368
Db      194  eTyrValValGlyGlnThrTrpGlnLysProTyrGlnGlnTyrTrpMetValAspCysTr 214
Qy      369  -----GAAGCAGCGGAGCGCATCTTGCACTTGTAGAAATAGATGCAAGCATC 417
Db      214  hrcysleuGlyGlnGlySerGlyArgIleThrCysTrnSerArgAsnArgCysAsnAspG 234
Qy      418  AGGACAACAAGACATCTTATAGAAATTGAGACACCTGAGACCAAGAAGATATCGAGAA 477
Db      234  InAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyA 254
Qy      478  ACTGCTCCAGTGCATGTGACAGCAAGCAAGCGCCAGGAGAGTGGAAAGTGAAGGACCA 537
Db      254  snleuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlnTrpLysCysGlnArgHisTr 274
Qy      538  CTTCTGTGACAGACATCATGAGGGGATCTGGCCCTTCAACCGATGTCGATTTGCT--- 593
Db      274  hrcerValGlnThrThrSerSerGlySerGlyProPheThrAspValArgAlaValT 294
Qy      593  ----- 593
Db      294  yrcLInProGlnProHisProGlnProProTyrGlnHisCysValAlThrAspSerGlyV 314
Qy      594  -----GGACTGAGTGGCTGCTAGACCGTTCATCTGTGCAACAAGCAAGCAT 639
Db      314  alValTyrSerValGlyMetGlnTrpLeu-----LysThrGlnGlnLysAsnLysGlnM 331
Qy      640  TGGTTGTAGCGTTGCTGTGATCTGTTGAGGGGACGAATCAAGACATTAATGCTTTT 699

```

```

RESULT 3
US-10-995-561-622
; Sequence 622, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559

```

Db	331	etleuCythrCylleuLyl-	----	337	
Oy	700	TTGAATTCGATCTTAAACATCAACGACTGCTCATGAGGAAAGACAGACGAGCTTAACTC	759		
Db	338	-----	AbnLylValSer-	341	
Oy	760	CAAAATCAAAAACATTTCGTACTGATATGGCGCGATGTCACTPAACTTGAGAAAGCTG	819		
Db	341	-----	341		
Oy	820	ACTTACTAAAGCGATCTTCAAGAACATTTGATCGTCAACGACGACGACTACT	879		
Db	342	-----	CysGlnGlnuthrAlaValThnGlnThrTyroLylYamSerIamG	357	
Oy	880	TTGAG-----GTCAATTGATTTTTCAGACGATGCAACCATTAATGATGCAACGCGACAG	933		
Db	357	LylGlnProCysValleuProPheThr	-----	TyrIamGlnArgT	370
Oy	934	TCYACTTTGCT-----	-----	GACAAAGATGGTTCGGTAACTTCGCCGACCAAC	978
Db	370	hrPheTySerCyethrThrglnLylArgGlnAspLylIleuThrPtySerThrThrs	390		
Oy	979	CTGTCCAGAAATTTTGTCTAAGCGGACATGGCGGTTAAGACCATATPAAAGAAAACCA	1038		
Db	390	erAenTyrgln-----	393		
Oy	1039	TACAAACCAAGCGAAATCTGTGTGATGTGAATATCTGTACAGTTTATCTCCCTTAAACC	1098		
Db	394	--GlnAspGlnIleTySerPheCysThrAspHisThrVal-----	406		
Oy	1099	CTGATGACGATTTTCAAGACAGGCTCAAAAGATCTAAGCTTATGAAAACTAGCTATCG	1158		
Db	406	-----	406		
Oy	1159	GTGACACCATGACATCTCAGAAATTACTAGCTCAACGACAAAGCATTTTAAACAAAACC	1218		
Db	407	-----	LeuValGlnThrArgLyl-----	412	
Oy	1219	ACCGAGGCTATACGATTTATGACAGCTGACTCTCCATCGTCACTCATGACATGACATTT	1278		
Db	413	-----	GlyIamSerAbnGlnAlaLeuCyHis-----	421	
Oy	1279	TCCGTAGCATTTTACCAATGATCAAGACTTACTTACCGTGTAAAAATCGGAAACAG	1338		
Db	422	-----	PheProPheLeuTyAsnAsnHis-----	429	
Oy	1339	CTTATAGCATCAATPAAAAATCTGTGTCGATGGAATTAACCAACCTAGACT	1398		
Db	430	-----	AsnTyThrAspCysThrS	436	
Oy	1399	CTGAGAAATATTCGTCCTTAAAAAAGGGAAAAAGCCGTATGATCCCTTATCGCAGTC	1458		
Db	436	erGlnLyl-----	ArgArgAspA	442	
Oy	1459	ACTTGAAA--CTGTTCACCATCAAAATACGTTGATGTCATPACCAAC-----	G	1503	
Db	442	smetLylStrPtySerLylThrThrGlnAenTyAspArgLnlYerPheGlnPheCysP	462		
Oy	1504	AATTGCTPAAAAAGTGAAGAGCTCTTAAACAGCTAGGAAACGTAACCTTAGACTTCAGAGATT	1563		
Db	462	romerAlaIalHisGlnLylIleCysThrThrAsnGln-----	474		
Oy	1564	TATACGATCTCGTNGATTAAGGCTAAACTACTTCAACCAATCTCGATGCTTTTGGTATTA	1623		
Db	475	-----	GlyValM	477	
Oy	1624	TGCACTATACCTTAACTGAAAAAGTGAAGATTAACAGATGACCAACCGTATCATTA	1683		
Db	477	et---TyrArgLl eolYaspGlnTrpAerLysGlnHisAspMecLylYHisMetMetArgC	496		
Oy	1684	CCGTTTATATGGGCAAGGCAAGCCCAAGGAAGAAATGCTAAGCTTACCATTTAGCTGTGTG	1743		

```

Db      496 /sThCyValGlyLysn-----GlyArgG 504
Qy      1744 GCCAGCGCCAAACAGATTGTA-----CCATAGCTGAGAGAGTGTTCATCATGCTGCTG 1797
Db      504 lyltlrThrCysIleAlaIatYrSerGlnLeuAdgApdInCysIle-----ValAspA 522
Qy      1798 GGACCTTCCTATGTCGTGGAGAAACGTGGAGAAACCCCTACCAAGAGCTGATGATGGTAG 185
Db      522 splleThrYAsnValAsnAspIrrPheNlsYAsrGlnIsdIugLugLysIleLeuA 542
Qy      1858 ATTGACTTGGCTCGGAGAGAGCGGACCGGACCGCATCATCTTCTAGAAATAGATGCA 191
Db      542 enCysThrCysePheIlyGlnGlyAsGlyAAsrTrpLysCysAspProValAspGlnCyseG 562
Qy      1918 ACATCAGACACACAGACATCTTATAGAAATTTGAGACACCTGAGACAGAAAGATATC 1977
Db      562 lnsAspserIuThrGlnIlyrPheYrGlnIleGlyAspserTrpGlnuYsYrValHis- 581
Qy      1978 GAGGAAACCTGCTCCAGTGCATCTTCACACAGGAAAGCGCCAGAGAGAGAGAGTAGTGA 203
Db      582 --GlyValAsrGlyrGlnCysYrYrGlyAsrGlyIleGlyGlnuTrpHisCysGln- 600
Qy      2038 GGACACCTCTGTGACACACATGACAGGAGGAGCTGGCCCC 2078
Db      601 -----ProLeuGlnIuThrYrProserSerSerGlyPro 611

RESULT 4
US-10-995-561-629
; Sequence 629, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 629
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-629

Alignment Scores:
Pred. No.: 2,1e-40 Length: 984
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
DB: 6 Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-629 (1-984)
Qy      5 GACCATTCATGTTGTGTCTCAGAGTCGACAGCGTTTGGACGACGATGCGTTCAAGT--- 61
Db      122 AapCyStrhCysIleGlyAlaGlyArgGlyArgIleSerCyThrIleAlaAsnArgCys 141
Qy      62 -----TCGTCGCGATTCGGTGATTCATCTTCCTAACAGTAGGCAAC 106
Db      142 HisGlnGlyGlnInsertyr-LysIleGlyAspYrTrp-----ArgArpR 157
Qy      107 CCGCAGCGCTAGCGGGGTCCTCAAGACAGAGACAGCATCATGGGAAACCGTGGCAGAGA 166
Db      157 HisGlnuTr----- 160
Qy      167 CCAACGCTGCCGAGATCTCGATCCGCGAAATTAATACATGACTATAGAGAGACCA 226
Db      160 ----- 160
Qy      227 CAAGCGTTCCCTCTAGAA---ATAATTTGTTTAACTTTAAGAAAGAGATATACATCG 283

```



```

Qy 1744 GCCAGCGCAGACAGATTGTA-----CCCATACCTGAGAGTGTTCATGATGCTGCTG 1797
Db 504 lYgluTrpThrCySileAlaTyRserGlnLeuArgAspGlnCySile-----ValAspA 522
Qy 1798 GGACTTCCTATGTGTGCGGAGAAAGCGGAGAAAGCGGAGAAAGCGGAGAAAGCGGAGAA 1857
Db 522 spIeThrTyRAsnValAspThrPheHieTyRAsnValAspThrPheHieTyRAsnValAsp 542
Qy 1858 ATTGACTTTCCTGCGGAGAAAGCGGAGAAAGCGGAGAAAGCGGAGAAAGCGGAGAAAGCGG 1917
Db 542 sncYThrCySHeGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 562
Qy 1918 AGGATCAGACACAAAGACATCTATAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1977
Db 562 lAspSerGlnThrGlnThrPheTyRGlNleGlnAspSerTrpGlnTyRValHis- 581
Qy 1978 GAGGAAAGCTGCTCCAGTGCATCTGCACAGGACAGCGCGGAGAGAGTGGAGTGTGAGA 2037
Db 582 --GlyValArgTyRGlNcySlyTyRclYArgGlylIeGlnGlnTrpHieCySgln- 600
Qy 2038 GGCACACCTGTCGACAGACATCGAGGAGATCTGGCCCC 2078
Db 601 -----ProLeuGlnTrpTyRProSerSerGlnPro 611

RESULT 6
US-10-995-561-623
; Sequence 623, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 623
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-623

Alignment Scores:
Pred. No.: 2,486-40 Length: 2355
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-623 (1-2355)
Qy 5 GACCATTCATGTTGTCAGAGTGGACAGCGTTTTCGACAGACAGTCCGTTACAGT--- 61
Db 122 AspCysThrCySileAlaTyRserGlnLeuArgAspGlnCySile-----ValAspA 141
Qy 62 -----TCGCTCGGATCGGTATCTGATTCATCTGCTAATCAGTAAAGGCAACC 106
Db 142 HisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 157
Qy 107 CCGCAGACCTAGCCGGGTCTCTCAACGACAGAGACAGATCATCGCACCCGTGGCCAGGA 166
Db 157 ohIeGlnThr----- 160
Qy 167 CCGAAGCTGCCGAGATCTGATCCCGGAAATTAATACACTCATATAGGAGAGCA 226
Db 160 ----- 160
Qy 227 CAAGGTTTCCTCTAGAA--ATAATTTTGTTTAATTAAAGAGATATACATG 283

```

```

Db 161 -GlyGlyTyRMetLeuGlnCySValCySLeuGlyAsnGlyTyRglGln- TrpThr----- 178
Qy 284 TGCAAGCAACAAGATTGTACCCCATAGCTGAGAGTGTTCATGATGCTGAGACTT 343
Db 179 -----CySAspTrpIleAlaGlnTyRAspPheAspHieAlaGlnTyR 194
Qy 344 CTAATGTGTCGAGAAACGTGGA----- 368
Db 194 erTyRValValGlnGlnTrpGlnTyRProTyRGlNglYTrpMetMetValAspCyS 214
Qy 369 -----GAGGCAAGCGGAGCATCTGCACTTCTTGAATAATGATGACATG 417
Db 214 hrcYsLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 234
Qy 418 AGGACACAGAGACATCTATAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 477
Db 234 lAspThrArgThrSerTyRArglIeGlnAspThrTrpSerTyRAspAsnArgGly 254
Qy 478 ACCTGCTCCAGTGCATTCGACAGGCAAGCGCGGAGAGAGTGGAGTGGAGAGCA 537
Db 254 snLeuLeuGlnCySileCySThrclYAsnGlyArgGlyGlnTrpTyRcySglnTyRglnst 274
Qy 538 CTTCTGTCAGACCAATCGAGCGGATCTGGCCCTTCACCGATGTTCTGATTCCT----- 593
Db 274 hrcerValGlnTrpThrSerSerclYserGlyProPheThrAspValArgAlaVal 294
Qy 593 ----- 593
Db 294 yrglnProGlnProHieAspGlnProProTyRglnHieCySValThrAspSerGly 314
Qy 594 -----GGACGTGAGTGGCTGTCAGACCGGTCATCTGTCACACAGCCCAT 639
Db 314 alValTyRserValGlnMetGlnTrpLeu-----LysThrGlnGlyAsnTyRgln 331
Qy 640 TGGTGTTCAGCTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTT 699
Db 331 etLeuCySThrCySleuGly----- 337
Qy 700 TTGAATTCATCTAACATCAACGCTGCTCATGAGAGAAAGACAGACAGCTTAAGTC 759
Db 338 -----AsnGlyValSer- 341
Qy 760 CAAATCAAAAACATTTGCTACTGATGTCGCGGATGTCATATAACTTGAGAAACCTG 819
Db 341 ----- 341
Qy 820 ACTTACTAAGGCTATTCAGAAACAATGATTCGTTACAGTCAACAGTACAGACATTA 879
Db 342 -----CySglnGlnThrAlaValThrGlnTrpTyRglYglYAsnSerArg 357
Qy 880 TTGAG-----GTCATGATTTTTCAGAGCATGCAACCATTAATGATCGAAACGGCAAG 933
Db 357 lYgluProCySValLeuProPheThr-----TyRAsnGlyArg 370
Qy 934 TCTACTTTGCT-----GACAAAGTGTGGTAACTTGGCCAGCCAGAC 978
Db 370 hrcHeTyRserCySThrThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 390
Qy 979 CTGTCAAGAAATTTTGTCTAAGCGGACATGCGCGGTTAGACCATATAAAGAAACCA 1038
Db 390 erAsnTyRglu----- 393
Qy 1039 TACAAACCAAGGAAATCTGTGATGTGAAATATATCTGATTAATCTCCCTTAAC 1098
Db 394 --GlnAspGlnTyRtyRserPheCySThrAspHieThrVal----- 406
Qy 1099 CTGATACGATTTCAACAGGTCCTAAAGATCTAAGTATTAAGAAACATGACTATG 1158
Db 406 ----- 406
Qy 1159 GTGACACATCAGATCTCAAGATTACTAGTCAAGCAAAAGATTTTAAACAAAAC 1218
Db 407 -----LeuValGlnThrArgGly----- 412

```

```

QY 1219 ACCGAGGTATAGATTTATGAACGTGACCTCCATTCGCTACTGATGACATGATTT 1278
DB 413 -----GlyAsnSerAsnGlyAlaLeuCysHis----- 421
QY 1279 TCCGATGATTTTACCAATGATTCAGAGATTTACTACCGTGTAAATCGGAAACAG 1338
DB 422 -----PheProheLeuTyraAsnHis----- 429
QY 1339 CTTATAGATTCATATAAAATCTGCTGTAATGAAGAAATTAACAACCTGACCTGATCT 1398
DB 430 -----AsnTyThrAspCysHis 436
QY 1399 CTGAGAAATATTACGCTTTAAAAAAGGGAAAAGCGGTATGATCCCTTGTATCGCAGTC 1458
DB 436 eTgUGly-----ArgArgAsp 442
QY 1459 ACTTGAAA---CTGTTCACCATCAATATACGTTGATGTCGATCAAC---G 1503
DB 442 smMetLysTrpCysGlyThrGlnAsnTyraAspAlaAspGlnLysPheGlyPheCysAr 462
QY 1504 AATTGCTTAAAAATGACAGCTCTTAACAGCTAGCGAAGCTAACTTAGACTTCAAGATTT 1563
DB 462 rOmEtAlaAlaHisGlyGlnIleCysThrThrAsnGlu----- 474
QY 1564 TATAGATCCCTGCTATAGGCTAAAGCTACTACAAACAATCTGCATCTTTGTGATTA 1623
DB 475 -----GlyVal 477
QY 1624 TGGATATATACCTTAACGTGAAAAGTAGAGATTAATGACGATGACAAACCGTATGATTA 1683
DB 477 et--TyraGlyIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetAlaGc 496
QY 1684 CCGTTTATATGGGCAAGCGAAGCGAAGAGAAATGCTAGTCACTTAAGCTGTGTGTG 1743
DB 496 ysthCysValGlyAsn-----GlyArg 504
QY 1744 GCCAGCGCAACAGATGTA-----CCATAGCTGAGAAAGTCTTTGATCATGCTGCTG 1797
DB 504 lYgIltTrpThrCysIleAlaTySerGlnLeuArgAspGlnCysIle-----ValAspA 522
QY 1798 GACCTTCATGCTGTCGGAAGAAAGTGGAGAAAGCCCTACGAAAGCTGGATGATGATGTG 1857
DB 522 spIleThrTyraAsnValAsnAspThrPheHisLysArgHisGlyGlnGlyHisMetLeuA 542
QY 1858 ATTGACTTCTCCCTGGAGAAAGGCAAGCGAAGCGATCACTTGCACTTGTAGAAATAGATGA 1917
DB 542 snCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysG 562
QY 1918 AGCATCAGGACACAGAGACATCTATAGAAATTTGAGACACCTGGAGCAAGAGATTAATC 1977
DB 562 lAspSerGlnTrpGlyThrPheTyrglnIleGlyAspSerTrpGlnLysTyraHis- 581
QY 1978 GAGGAACCTGCTCCAGTGCATCTGACAGGCAAGCGAAGCGAGAGAGTGAAGTGGAGA 2037
DB 582 --GlyValArgTyrglnCysTyrglyArgGly11eGlyGlnTrpHisCysGln- 600
QY 2038 GGCACACCTCTGTGACAGACACATGACGAGGATCTGGCCCC 2078
DB 601 -----ProLeuGlnThrTyProSerSerSerGlyPro 611

```

```

; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-627

Alignment Scores:
Pred. No.: 2,48e-40 Length: 2355
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
DB: 6 Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-627 (1-2355)
QY 5 GACCATTCATGTTGTTGTCAGGTCGACAGGTTTTCAGACAGAGATCGCTTCAAGT--- 61
DB 122 AspCysThrCysAlleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCys 141
QY 62 -----TCGCTCGGCTATCGGTGATTCATCTCGTAAACAGTAAGCAACC 106
DB 142 HisGlyGlyGlyGlnSerTyrgLysIleGlyAspThrTrp-----ArgArgPr 157
QY 107 CCGCCAGCTAGCGCGGGTCTTCAACAGACAGACAGATGATGCGCACCGGTGGCAAGA 166
DB 157 oHisGluThr----- 160
QY 167 CCGAAGCTCCCGAGATCTGCATCCCGGGAATTAATAGACTCACTATAGGGAGACA 226
DB 160 ----- 160
QY 227 CAACGTTTCCCTTAGAA---ATAATTTGTTTACTTTAAGAGAGATATATCCATGG 283
DB 161 -GlyGlyTyrgMetLeuGlnCysValCysLeuGlyAsnGlyLysGlyLys-TrpThr----- 178
QY 284 TGCAGACACAGATTTGTAACCAATAGCTGAGAAAGTCTTTGATCATGCTGCTGGAATT 343
DB 179 -----CysLysProIleAlaGlnLysCysPheAspHisAlaAlaGlyHis 194
QY 344 CCTATGTCGTCGGAAGAAAGCTGGA----- 368
DB 194 eTyrgValValGlyGlnTrpGlnLysProTyrglnGlyTyrmMetValAspCysT 214
QY 369 -----GAAGCAGCGGAGCGATCACTTGCACTTCTAGAAATAGATGCAACGATC 417
DB 214 hrcysleuGlyGlyGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnArg 234
QY 418 AGGACACAGAGACATCTTATGAAATTTGAGACACCTTGAGCAAGAGATATATGAGAA 477
DB 234 lAspThrArgTrpHiserTyrgIleGlyAspThrTrpSerLysAspAsnArgGlyA 254
QY 478 ACCTGCTCCAGTGCATCTGACAGGCAAGCGGAGAGAGTGAAGTGAAGGAGACA 537
DB 254 snLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGlnTrpLysCysGlyLysHis 274
QY 538 CCTGTGACAGACACATGACAGGAGATCTGACCGCTTCAACGATGTTGATTTGCT----- 593
DB 274 hSerValGlnThrHiserSerGlySerGlyProPheHiserValArgAlaAlaValT 294
QY 593 ----- 593
DB 294 yrglnProGlnProHiserGlnProProProTyrglnHisCysValThrAspSerGlyV 314
QY 594 -----GACCTGAGTGGTCTGTAACGCTCCATCTGTCAACAACGCAAT 639
DB 314 aValTyrgSerValGlyMetGlnTrpLeu-----LysThrGlnGlyAsnLysGlnM 331
QY 640 TGGTGTGATGCGTGTGCTGACTGTTGAGGGGAGCAATGAACATTAAGTCTTAATTTT 699
DB 331 etLeuCysThrCysLeuGly----- 337

```

OY	700	TTGAANTGATCTAATCATCAGACCTGCTCATGAGAGAAAAGACAGACGAGCTTAAGTC	759
Db	338	-----AsnGlyValSer-----	341
OY	760	CAAAATCAAAACCATTTGCTACTGATAGTGGCGGATGTCACATTAACCTTGAGAAACGTG	819
Db	341	-----	341
OY	820	ACTTACTAAAGGCTATTCAAGACAATTGATCGCTAACGTCACAGTAACGACGACTACT	879
Db	342	-----CysGlnGluThrIleAlaValThrGlnThrTyrCylGlyAsnSerAaG	357
OY	880	TTGAG-----GTCAATGATTTTGGACGGATGCAACATTACTGATGAAACGGCAAG	933
Db	357	LyGluProCysValLeuProPheThr-----TyrAsnGlyArgT	370
OY	934	TCTACTTGGCT-----GACAAAGATGGTTCCGTAACCTTGGCCAGCCCAAC	978
Db	370	hrPheTyrSerCysThrThrGluGlyArgGlnAspGlyHisLeuTyrCysSerThrTrpS	390
OY	979	CTGTCCAGAAATTTTGGCTAAAGCGACATGTGCGGCTTAAGACCATATTAAGAAAAACCA	1038
Db	390	erAsnTyrGlu-----	393
OY	1039	TACAAACACCAAGCCAAATCTGTGTAATGTGGAATATATCTGTACAGTTTACTCCCTTAAC	1098
Db	394	--GlnAspGlnLysTyrTyrSerPheCysThrAspHisThrVal-----	406
OY	1099	CTGATGACGATTTTCAGACCAAGGTCCAAAGATACTAAGCTATTGAAAAACATAGCTATCG	1158
Db	406	-----	406
OY	1159	GTGACACCATGCATCTCAAGAAATTACTAGCTCAAGACAAAGCATTTTAAACAAAACC	1218
Db	407	-----LeuValGlnThrArgGly-----	412
OY	1219	ACCAAGCTTAACGATTTATGAACGTGACTCTCCATCTGCACTTCATGACAAATGACATTT	1278
Db	413	-----GlyAsnSerAsnGlyAlaLeuCysHis-----	421
OY	1279	TCCGTAAGATTTTACCAATGATGCAAGGTTTACTTACCCGGTTAAAAATCGGAAACAAG	1338
Db	422	-----PheProPheLeuTyrAsnAsnHis-----	429
OY	1339	CTTATAGGATCAATAAAAAATCTGGTCTGAATGAAGAAATATCAACAACCTGACTACT	1398
Db	430	-----AsnTyrThrAspCysTrpS-----	436
OY	1399	CTGAGAAATATTACGTCCTTAAAAAAGGGAAAAAGCCGTATGATCCCTTTGATCGCAATC	1458
Db	436	erGluGly-----ArgTrpAspA-----	442
OY	1459	ACTTGAAA--CTGTTCACCATCAAAATACGTGATGTGATACCAAC-----G	1503
Db	442	smMetLysTyrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysP	462
OY	1504	AATGCTAAAAAGTGGAGCGCTTTAAGCAGTAGGAAACGTAACTTAGACTTCAAGATTT	1563
Db	462	rometAlaIleHisGluGluIleCysThrThrAsnGlu-----	474
OY	1564	TATACGATCTCGTGAATAGGCTAAACTACTCTACACAACTCGATGCTTTTGATTA	1623
Db	475	-----GlyValM-----	477
OY	1624	TGACATATCCTTAATCTGAAAAAGTAGAGATATACAGATGACCAACCGTATCAATA	1683
Db	477	et--TyrArgIleGlyAspGlnTyrAspLysGlnHisAspMetGlyHisMetMetArgC	496
OY	1684	CCGTTTATATGAGGACGACCGCAAGAGAGAAATGCAAGTACCATTTTAGCTGCTGCTG	1743
Db	496	ythrCysValGlyAsn-----GlyArgG-----	504

```

QY 1744 GCCAGGGCGCAACAGATTGTA-----CCCATACGCGAGAAATGTTTGGATCATGCTCTGCTG 1797
Db 504 1yglutprTh-Cys11lealatySerglnleuAtgAspGlnCys1le-----ValAspA 522
QY 1798 GGACTTCCTATATGCTGCGGAGAAAGTGGAGAAAGCCCTTACCAAGGCGCTGATGATGCTAG 1857
Db 522 ap1lethrThyAsnValAsnAspRhrPheH1slybAArgH1sG1uGlnG1yH1sMetLeuA 542
QY 1858 ATTGTACTTTCCTGGGAGAGAGCGACGGAGCGCATCATCTGCACTTCTTAAGAAATGATGCA 1917
Db 542 snCysThrCysPheGdlYglnlYatrgYlYatrgTPrlyCyAspAspProValAspGlnCysG 562
QY 1918 ACGATCAGACAAACAAGACATCCATATGAAATTGAGACACCTGGAGCAAGAAGATATATC 1977
Db 562 1nAspSerGlnHnG1YThrPheTrglnl1eG1yAspSerTrpGlnYerThyValH1s- 581
QY 1978 GAGGAACCTGCTCCATGCATGCATCTGCAACAAGCGCCGAGAGAGTGAAGTGTAGA 2037
Db 582 --GlyValArgTyGlnClyerThyCysThyrg1YatrgG1y11eG1YglnTrpH1sCysGln- 600
QY 2038 GGCACACCTGTGTGCAAGCCACATCAGACGGGATCTGGCCCC 2078
Db 601 -----ProleuGlnThrTyRProSerSerSerg1yPro 611

RESULT 8
US-10-821-234-1545
; Sequence 1545, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC_SEQ_genes Version 1.0
; SEQ ID NO 1545
; LENGTH: 2384
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-821-234-1545

Alignment Scores:
Pred. No.: 2,48e-40 Length: 2384
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
Gaps: 28
DB: 6

US-09-940-235-12 (1-2096) x US-10-821-234-1545 (1-2384)
QY 5 GACATTCATGTTTGTGCTCAGGTCCGAGACGTTTGCACAGACAGATCGCTTACGT--- 61
Db 151 AspCysThrCys11leG1yAlaG1yArg1YAsp11eSerCysThr11leAlaAsnArgCys 170
QY 62 -----TCGCTCGGCTATCGGATGATTCATTCGCTAACATAGAGCAACC 106
Db 171 H1sGlnG1yAlGlnSertYr-Lys11eG1yAspHnTrp-----ArgArgPr 186
QY 107 CCGCAGACCTAGCCGGGCTCTCAACAGACAGACAGATCATGCGACCCGTGGCAGAGA 166
Db 186 oh1sGlnThr----- 189
QY 167 CCGAAGCTGCCGAGATTCGATCCCGCGCAAAATTAACGATCACTATAGGAGACCA 226
Db 189 ----- 189

```



```

CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FaSTSeq for Windows Version 4.0
SEQ ID NO 626
LENGTH: 2386
TYPE: PRF
ORGANISM: Homo sapiens
US-10-995-561-626

Alignment Scores:
Pred. No.: 2,48e-40 Length: 2386
Score: 634.50 Matches: 207
Percent Similarity: 36.69% Conservative: 70
Best Local Similarity: 27.42% Mismatches: 151
Query Match: 17.03% Indels: 328
DB: Gaps: 28

US-09-940-235-12 (1-2096) x US-10-995-561-626 (1-2386)

OY 5 GACCATTCAATGTTGTTGCTCAGGTCGCGACGCTTTTGACAGCAGATCGCTTCACTG--- 61
Db 122 AapCystrhCysrleeglyalaglyArgilAeserCystrhrllealAasrArgCys 141
OY 62 -----TCGCTCGGGTATTCGCTGATTCATTTCTGCTAACAGTAGGCAACC 106
Db 142 HhsglmglyglinserrYr-lysrleeglyaapThrrp-----ArgargPr 157
OY 107 CGCGCAGCTTAGCGGGTCTCAACGACAGAGCAGATCATGCGCACCGTGGCCAGGA 166
Db 157 ohlbglnthr----- 166
OY 167 CCACACGCTGCCGAGATCTCATGCCGCGAAATTAAACACTCACTATAGGAGACA 226
Db 160 ----- 160
OY 227 CAACGGTTCCCTCTAGAA---ATAATTTGTTAACTTTAAGAGAGATATACATGG 283
Db 161 -gllylYrmetleuglucyvalAcyseuglyAasnglyLyglYglu-TrrpThr--- 178
OY 284 TGCAGACACAACAGATGTATACCATATGACGAGAGAGTTTGAATCATGCTGCGGACCT 343
Db 179 -----CylysrprrllealaglnlyscyspneaprrlsalalaglyThrs 194
OY 344 CCTATGTGTCGAGAGAACTGGGA----- 368
Db 194 erlyrValValglYglutrrtrpnlulYerprolyrnglnlyrTpmetmetValaspcystr 214
OY 369 -----GAGGACGAGGACGATCATCTTGACCTTTAGAAATATAGTCAAGATC 417
Db 214 hrCyseuenglYglunglyserglYarglIethrCystrhrserArGAsrArgCysasnsaprg 234
OY 418 AGGACACAAGGACATCTCTATGAATTGGAGACACTGGAGCAGAAGAGATATACGAGAA 477
Db 234 lnsprhrtrgtrhrserlyrArglleglyAsprhrtrpserlyslYsrAspAsrArglYA 254
OY 478 ACCTGCTCAGTGCATCTGCACACAGGACAGCGCCGAGAGAGAGTGAAGTGAAGGACCA 537
Db 254 snleuenglncysrlleCystrhnglyAasnglyArgglYglutrrpYsrCyglulnrglhr 274
OY 538 CCTGTGTGACAGACATGAGCGGATCTGGCCCTTACCCGATGTTCGTAATGCT--- 593
Db 274 hrserValglnthrrhrserserglYserclYrProphetrhrasrValarglAlaValr 294
OY 593 ----- 593
Db 294 yrGlnprrGlnProhlsProGlnProProProlyrclYhlsCyseValThraspserrclYv 314
OY 594 -----GGACCTGAGGTGGCTGTAAGCCTGCATCTGTGCAACACAGCCAT 639
Db 314 alVallyrserValglYmetGelntrrpHeu-----lystrhnglnlyAasnlYseln 331
OY 640 TGGTGTATTAGCGTGTGCTGTATCTGTAGAGGAGCAATCAAGCATTAATGCTTAAATTTT 699

```

Db	331 :::		etLeuCyserthrCyLeuGly	337
Oy	700	TTGAATGCATCTAATCATCAAGACGCTGCTATGAGAAAGACAGACAGGCTTAAGTC			759
Db	338	-----		AsnIleValSer	341
Oy	760	CAAAATCAAAACCAATTGCTACTGATAGTGGCGCGATGTCACTATAACTTGAGAAAGCTG			819
Db	341	-----			341
Oy	820	ACTTACTAAAGCTATTGCAAGAACATGATGCGCTAAGCTCAGACAGTAAACGACACTACT			879
Db	342	-----			357
Oy	880	TTGAG-----GTCAATGATTTTGGCAAGCAGATGCAACCACTTATGATGGAACCGCAGG			933
Db	357	LyGIuProCyValIleuProPheThr		LyTrAsnIleValGyt	370
Oy	934	TCCTACTTGTCT-----GACAAAGATGTTGCGTAACTTTGCGGACCCCAAC			978
Db	370	hrPheTySerCyserThrThrGluGlyAsgIAsnArgIyHniIleuThrCyserThrThrTs			390
Oy	979	CTGTCCAGAAATTTTGTCTAAGCGGACATGTGCGGTTAGACATATATAAGAAAAACCA			1038
Db	390	erAsnTyGly-----			393
Oy	1039	TACAAACCAAGCAAAATCTGTGATGTGAATATACTGTACATTTACTCCCTTAAAC			1098
Db	394	--GlnAsnGlnIlyIyIserPheCyThrIAsnPhIsthrVal-----			406
Oy	1099	CTGATGACGAATTTGACACGAGGTCTCAAAAGATACTAAGCTATTGAAAACTAGCTATCG			1158
Db	406	-----			406
Oy	1159	GTGACACATCATCATCTCAAGATTAATCTGCTCAAGCACAAGCAATTTTAAACAAAAC			1218
Db	407	-----		LeuValGlnThrArgGly	412
Oy	1219	ACCAGGCTATACGATTTATGACGCTGCTCCTCAATCGTCACTGACATGACAAATGACATTT			1278
Db	413	-----		GlyAsnSerAsnGlyAlaLeuCyHniS	421
Oy	1279	TCCGTACGATTTTACCAATGATCAGAGATTTACTACCGGTTAAAAATCGGAAACAG			1338
Db	422	-----		PheProPheLeuTyTrAsnAsnHniS	429
Oy	1339	CTTATAGCATCAATTAATAAATCTGGTCTGAATGAAGAAATTAACACACTGACTGATCT			1398
Db	430	-----		AsnTyTrThrAspCyThrTs	436
Oy	1399	CTGAGAAATATTAGCTCTTAAAAAGGGGAAAAAGCGTATGATCCCTTGGATGCGCAGTC			1458
Db	436	erGluGly-----		ArgTrAsnArg	442
Oy	1459	ACTTGAAA-----CTGTTCACATCAATATACCTGTGATGTCATACCAAC-----G			1503
Db	442	smMetIyGtrCyseGlyThrThrGlnAsnTyIAsnArgIlyAsnArgIlyAsnArgIlyPheCyse			462
Oy	1504	AATGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCAGGACGTAACCTTAGACTTACAGAGATT			1563
Db	462	rometAlaIAsnIleGlnIlyIleCyserThrThrAsnGly-----			474
Oy	1564	TATAGCATCCCGTGATAGGCTAAAGCTACCTACCTACAAATCTGATGCTTTGGTATTA			1623
Db	475	-----		GlyValM	477
Oy	1624	TGACATATACCTTAAGTGAAGAGTAAAGATATACAGATGACACCAACCGTATTA			1683
Db	477	et---TyArgIleGlyAsnArgIlyTrAsnArgIlyAsnArgIlyIleAsnMetIyHniIleMetMetArg			496
Oy	1684	CCGTTTATATGCGGACGACCGCAAGAGAGAAATGCTAGCTACATTTAGCTGGTGTG			1743

```

Db      496 ysthrCyseValglYasn-----GLYArg 504
Qy      1744 GCCAGGCCGCAACAGATTGTA-----CCCATAGCTGAGAAAGTCTTTGATCATGCTGCTG 1797
Db      504 lYglutRprThrCysllealalyrSerglnLeuArgspglnCyslle-----ValAspA 522
Qy      1798 GCAGTCTCTATGTCGTGCGAAGAACCTGGAGAAAGCCCTACCAAGCTGGATGATGATG 1857
Db      522 spllleThrTyraValaValaAspThrPheHileValArgHilegluGluGlyHileMetLeuA 542
Qy      1858 ATTGTACTTCTGCTGGAGAGGACGAGCGCATCTGCACTTCTAGAAATAGATGCA 1917
Db      542 snCystrCyseHegleGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCys 562
Qy      1918 ACAGTACAGACACAGACATCTTATAGATTGGAGACCTGGAGCAAGAGATATATC 1977
Db      562 lnaSpSerGlnThrlyThrPheTyrglnIleGlyAspSerTrpGluTyryValHis- 581
Qy      1978 GAGCAACCTGCTCGATGTCATCTGACAGGACGCGGAGAGAGAGTGTGAGA 2037
Db      582 --GlyValArgTyrglnCysTyryGlyArgGlylleglyGluTrpHisCysGln- 600
Qy      2038 GGCACACCTCTGTGACACCATGAGCGGATGCGCC 2078
Db      601 -----ProleuGlnThrTyryProSerSerSerdlyPro 611

```

```

RESULT 10
US-10-631-558-4
; Sequence 4, Application US/10631558
; Publication No. US20050260598A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-631-558-4

```

```

Alignment Scores:
Pred. No.: 2,5e-40 Length: 259
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 16.96% Indels: 6
DB: Gaps: 3

```

```

US-09-940-235-12 (1-2096) x US-10-631-558-4 (1-259)
Qy      1698 AAGCGACCCGAGAGAGAAATGCTAGTACCATTTA-----GCTGGTGTGTC 1745
Db      124 ArgArgProHis---GluThrGlyGlyTyryMetLeuGluCysValCysleuGlyAsnGly 142

```

```

Qy      1746 CAGCGCGAA---CAGATTGTACCATATGATGAGAAAGTCTTATGATCATGCTGAGACT 1802
Db      143 lYsGlyGluTrpThrCysLysProIlealGlnLysCysPheAspHisAlaIleGlyThr 162
Qy      1803 TCCTATGTGTGTCGAGAAAGCTGGAGAAAGCCCTACCAAGCTGGATGATGATGATG 1862
Db      163 SerTyValValGlyGluTrpThrTrpGluTyryProTyrglnGlyTrpMetValAspCys 182
Qy      1863 ACTTGGCTGGAGAGAGGACGAGCGGATCATCTTGCATCTTCTAGAAATAGATGCAACGAT 1922
Db      183 ThCyseuGlyGlyGluYserGlyArgIleThrCystrSerTrpAspAsnAsp 202
Qy      1923 CAGGACACAGGACATCTTATAGATTGAGACACCTGGAGCAAGAGATATCGAGA 1982
Db      203 GlnAspThrArgThrSerTyryArgIleGlyAspThrTrpSerTyryAspAsnArgGly 222
Qy      1983 AACCTCTCAGATGATCTTGTACAGGCGAAGCGCGAGAGAGTGAAGTGAAGAGCAC 2042
Db      223 AsnLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
Qy      2043 ACCTTGTGACACCATGTCAGACGAGATGCGCCCTTACCCAGATGTTGCT 2093
Db      243 ThrSerValGlnThrThrSerSerGlySerGlyProPheThrAspValArg 259

```

```

RESULT 11
US-11-006-119-31
; Sequence 31, Application US/11006119
; Publication No. US20050260691A1
; GENERAL INFORMATION:
; APPLICANT: Ndao, Momar
; APPLICANT: Ward, Brian
; APPLICANT: Caffrey, Rebecca
; APPLICANT: Spithill, Terry
; APPLICANT: Li, Hongshan
; APPLICANT: Podust, Vladimir
; APPLICANT: Perichon, Regis
; APPLICANT: CIPHERGEN Biosystems, Inc.
; TITLE OF INVENTION: Serum Biomarkers for Chagas Disease
; FILE REFERENCE: 016866-012130US
; CURRENT APPLICATION NUMBER: US/11/006,119
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/527,153
; PRIOR FILING DATE: 2003-12-05
; PRIOR APPLICATION NUMBER: US 60/565,093
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US 60/625,519
; PRIOR FILING DATE: 2004-11-06
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: N-terminal 28.7 fragment of Fibronectin
; US-11-006-119-31

```

```

Alignment Scores:
Pred. No.: 2,57e-40 Length: 300
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 16.96% Indels: 6
DB: Gaps: 3

```

```

US-09-940-235-12 (1-2096) x US-11-006-119-31 (1-300)
Qy      1698 AAGCGACCCGAGAGAGAAATGCTAGTACCATTTA-----GCTGGTGTGTC 1745
Db      124 ArgArgProHis---GluThrGlyGlyTyryMetLeuGluCysValCysleuGlyAsnGly 142
Qy      1746 CAGCGCGAA---CAGATTGTACCATATGATGAGAAAGTCTTATGATCATGCTGCGGACT 1802

```



```

Db      143  LysGlyGluThrCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 162
Qy      1803  TCTATGATGATCGAGAAAGCTGGAGAAAGCCCTACCAAGGCTGATGATGATGATGAT 1862
Db      163  SerTyValValGlyGluThrTrpGluLysProIleGlnGlyTrpMetCysValAspCys 182
Qy      1863  ACTTCCTGGAGAGAGCGAGCGAGCCATCATTGCACTTTAGAAAATAGATGCAACGAT 1922
Db      183  ThrCysLeuGlyGluGlySerGlyArgGlyLeThrCysThrSerArgAsnAlaGlyAsnAsp 202
Qy      1923  CAGGACACAAAGACATCTTATAGATTTGAGACACCTCGAGCAAGAGATATCGAGA 1982
Db      203  GlnAspThrArgThrSerTyArgGlyLeGlyAspThrTrpSerTyLysAspAsnArgGly 222
Qy      1983  AACCTGCTCAGATGATCTGCACAGAGCAAGCGCGAGAGATGTGAGAGAGGAC 2042
Db      223  AsnLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 242
Qy      2043  ACCCTGTGCAGACCAATCGATCGAGGATTCGCCCTTCACCGATTTCT 2093
Db      243  ThrSerValGlnThrThrSerSerGlyProPheThrAspValArg 259

RESULT 12
US-10-995-561-631
; Sequence 631, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; LENGTH: 642
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-631

Alignment Scores:
Pred. No.: 2,978-40 Length: 642
Score: 632.00 Matches: 117
Percent Similarity: 87.59% Conservative: 3
Best Local Similarity: 85.40% Mismatches: 11
Query Match: 16.96% Indels: 6
DB: Gaps: 3

US-09-940-235-12 (1-2096) x US-10-995-561-631 (1-642)
Qy      1698  AAGCGACCCGAGAGAGATGCTAGCTACATTTA-----GCTGGTGTGC 1745
Db      155  ArgGluProHis---GluThrGlyGlyTyMetLeuGluCysValCysLeuGlyAsnGly 173
Qy      1746  CAGGCGCAA--CAGATTGATCCATAGCTAGAGAGCTGTTTATCATGCTGCTGAGCT 1802
Db      174  LysGlyGluThrTrpCysLysProIleAlaGluLysCysPheAspHisAlaAlaGlyThr 193
Qy      1803  TCTATGATGATCGAGAAAGCTGGAGAAAGCCCTACCAAGGCTGATGATGATGATGAT 1862
Db      194  SerTyValValGlyGluThrTrpGluLysProIleGlnGlyTrpMetCysValAspCys 213
Qy      1863  ACTTCCTGGAGAGAGCGAGCGAGCCATCATTGCACTTTAGAAAATAGATGCAACGAT 1922
Db      214  ThrCysLeuGlyGluGlySerGlyArgGlyLeThrCysThrSerArgAsnAlaGlyAsnAsp 233
Qy      1923  CAGGACACAAAGACATCTTATAGATTTGAGACACCTCGAGCAAGAGATATCGAGA 1982
Db      234  GlnAspThrArgThrSerTyArgGlyLeGlyAspThrTrpSerTyLysAspAsnArgGly 253
Qy      1983  AACCTGCTCAGATGATCTGCACAGAGCAAGCGCGAGAGATGTGAGAGAGGAC 2042

```

```

Db      254  AsnLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 273
Qy      2043  ACCCTGTGCAGACCAATCGAGGATTCGCCCTTCACCGATTTCT 2093
Db      274  ThrSerValGlnThrThrSerSerGlyProPheThrAspValArg 290

RESULT 13
US-10-995-561-621
; Sequence 621, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 1341
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-621

Alignment Scores:
Pred. No.: 1,028-13 Length: 1341
Score: 278.50 Matches: 136
Percent Similarity: 34.32% Conservative: 73
Best Local Similarity: 22.33% Mismatches: 195
Query Match: 7.47% Indels: 205
DB: Gaps: 25

US-09-940-235-12 (1-2096) x US-10-995-561-621 (1-1341)
Qy      397  CTAGAATAAGATGCAAGATCAGACACAAAGACATCTTATGAAATTGAGACACTGGA 456
Db      784  MetGluSerGlnProIleGlyThrGlnSerThrAlaIle-----ProAlaProThr 801
Qy      457  GCAAGAGATTAATCGAGAAAGCTGCTCGATGATCTGCAAGGACGCGAGAGAG 516
Db      802  AsnLeuLysPheThrGlnValThrProThrSerLeuSerAlaGlnTrpHis-ProPheAs 821
Qy      517  AGTGAAGTGTGAGGACACACCTCTGTGACACACATCGAGGAGGATCTGCGCTTCA 576
Db      821  nValGlnLeu-ThrGlyTyArgValArgValThrProLysGluLysThrGlyProMetL 841
Qy      577  CCGATGTCGATTTGCTGACCTGAGCTGAGTGGCTGTAACCGTCACTTGTCAACAAGCC 636
Db      841  LysGluIleAsnLeuAla-----ProAspSerIle 851
Qy      637  AATTGCTTTAGGCTGTGCTAGCTGTGAGGAGCAAGATCAGACATTAGCTTAAT 696
Db      851  erValValIleSer-----GlyLeuMetValAlaThrLysTyGlnValSerValTyra 869
Qy      697  TTTTGAATGCAATTAACATCAGACCTGCTCAGAGAAAGACAGAG--CAAGGCT 753
Db      869  IaleuLysAspThrLeuThrSerArgProAlaGlnGlyValIleThrThrLeuGluLys 889
Qy      754  TAACTCCAAATCAAAAACATTGCTATGATGAGCGGAGATGCAATTAACCTTGA 813
Db      889  alSerProProAlaGlyAlaGlyValThrAsp----- 899
Qy      814  AAGTGACTTAATTAAGGCTATTCAAGAAACATTGATCGCTAAGTCCACAGTAACGAG 873
Db      899  ----- 899
Qy      874  ACTACTTGAAGTCATGATTTTGCAGCGATGCAACATTACTATCGAAACGCGAAG 933
Db      900  -----AlaThrGluThrThrIleThrIle----- 907

```

```

Qy 934 TCTACTTGCTGCAAAAGATGGTTCGTGAACCTTGGCAACCAACCTGCTCAAAATTTT 993
Db 908 -----SerTrpArgThrIleuThrIleThrglyPheg 920
Qy 994 TGCCTAACGCGACATGCGCGTTAGACCATTAAGAAAAACCAATACAAACCAAGCA 1053
Db 920 InValAspAla-----ValProAlaasnIyIlnThrProIleGlnArgThrIle 937
Qy 1054 AATCGTGTAGTGGAAATATACGTACATTTACTCCCTTAAACCTGATGAGATTTCA 1113
Db 937 yspProaspValArgSerTrpIlnIle-----ThrgIyLeuGlnProGlyThrAspIyrl 955
Qy 1114 GACCAAGTCTCAAGATTAAGTAAAGTAAAGCACTAGCTATCGGTGACACCATCAT 1173
Db 955 ys-----IleTyrlleuTyrlrIleuAsnAspAsnAlaArgS 967
Qy 1174 CTCAGAAATTAAGTCAAGCA----- 1196
Db 967 erSerProValValIleAspAlaSerThrAlaIleAspAlaProSerAsnIleuArgPheL 987
Qy 1197 -----CAAGCATTTTAAACAAACCAACCA-----G 1224
Db 987 euAlaIleThrProAsnSerIleuValSerTrpIlnProProArgAlaArgIleThrg 1007
Qy 1225 GCTATACGATTT--TATGAACGTGACCTCTCAATCGTCACTGATGACAAATGACATTTTCC 1281
Db 1007 IYTyrlleIleIySerTyrluIySProGlySerProPro-----A 1020
Qy 1282 GTACGATTTTACCAATGATCA----- 1304
Db 1020 rglValValIleProArgProArgProGlyValIlnThrgIlnAlaThrIleThrgIyLeuGlnP 1040
Qy 1305 -----GAGTTTACTTACCGTGT-----AAAAATGGGGAACAAGCTTATAGGA 1347
Db 1040 roGlyThrgIyThrgIyThrgIleTyrlIleAlaIleuIyAsnIlnGlnIySerGlnProl 1060
Qy 1348 TC-----AATAAAATCGTGTGTAATGAAGAAATTAACAACACTGACCTGATCTCG 1401
Db 1060 euIleGlyArgIyIySthrgIy--GlnGlnAlaIleuSerGlnIlnThrIle----- 1076
Qy 1402 AGAAATATTACGCTCTTAAAAAAGGGGAAAGCCGATGATCCCTTGAATCGCAGTCACT 1461
Db 1077 -----SerTrpIlnProPheGln----- 1082
Qy 1462 TGAACCTGTCACCATCAATACGTTGATGTGATCCGACCAAGATTCCTAAAAAGTAGC 1521
Db 1083 -----AspThrSerGlnTyrlIleIleSerGlySth 1092
Qy 1522 AGCTCTTAACAGCTAGCGAAGCTTAAGCTTGAAGATTTATGATTCCTCGATGA 1581
Db 1092 IspProValGlyThrAspGlnGlnProLeuGlnPheArg-----ValProGlyThrs 1109
Qy 1582 AGGCTAAACTACTTACAAACAATCTCGATGCTTTGGTATTATGACTATTAACCTG 1641
Db 1109 erThrSerAla-----ThrIleuThrg 1116
Qy 1642 GA-----AAAGTAGAGCATTAATC 1659
Db 1116 IyLeuThrArgIyAlaThrTyrlAsnIleIleValGlnAlaIleuIyAspGlnGlnArgH 1136
Qy 1660 AGGATGACCAACCGATCATCAACCGTTATATATGCGCAAGGACCCGAGAGAGAGAA 1719
Db 1136 IlySValArgIyGlnGlnValIlnValIlnValGlyAsnSerValAsn--GlnGlyLeuAsnG 1155
Qy 1720 CTAGCTACCATTTAGCTGTGTGTGCGCAAGCGCAACAGATTGATCCCATAGCTGAGAA 1779
Db 1155 In-----ProThrAspAspSerC 1161
Qy 1780 GTTTGATCATGCTGCTGGACTTCTATGTGTGCGAGAAACGTGGAGAAAGCCCTTACC 1839
Db 1161 yspPheAspProTyrlrThrValSerHsIyAlaValGlyAspGlnIlnThrgIlnArgMetSerC 1181
Qy 1840 AAGCGTGATGATGTGAATTCTACTTGTGCTGGAGAAAGCGGACCCATCATCTTGCA 1899

```

```

Db 1181 IuSerGlyPheIySleuIleuCyGlnIySleuGlyPheGlySerGlyHsIlePheArgCySa 1201
Qy 1900 CTCTTAGAAATTAAGTCAACAGATCAAGACACAAAGACATCTTATGAATTTGAGACACT 1959
Db 1201 spSerSerArgTrpCySthIleAspAsnGly-----ValAsnTyrlIySileGlyGlnIySth 1219
Qy 1960 GGAGCAAGAGATTAATTCAGAGAAACCTGCTCCAGTGCATCTGACAGGCAAGCGCGAG 2019
Db 1219 rPaspArgGlnGlyGlnIyAsnGlyGlnIyMetSerSerTrpCySleuGlyAsnIyIySg 1239
Qy 2020 GAGAGTGAAGTGTGAGAGGAC 2042
Db 1239 IyGlnPheIySlyAspProHs 1246

RESULT 14
US-10-995-561-632
; Sequence 632, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 632
; LENGTH: 693
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-632

Alignment Scores:
Pred. No.: 1,51e-13 Length: 693
Score: 275.50 Matches: 142
Percent Similarity: 32.88% Conservative: 77
Best Local Similarity: 21.32% Mismatches: 220
Query Match: 7.39% Gaps: 25
DB: 6

US-09-940-235-12 (1-2096) x US-10-995-561-632 (1-693)
Qy 346 TATGTGCTGAGAAACGTGGAGAAAGCGAGCGATCACTTGCACTTTAGAAATA 405
Db 61 TyrlValIleSerValTyrlAlaGlnAsnProSerGlyIySleuGlnProLeuValGlnThr 80
Qy 406 GATGCAACGATCAAGACACAAAGACATCTTATGAATTTGAGACACTGAGCAAGAGG 465
Db 81 AlaValThr-----ThrIleProAlaProThrAspLeuIyS 92
Qy 466 ATATGAGGAACCTGCTCAGATGATTCAGACAGGACCGCGAGAGTGAAGT 525
Db 93 PheThrGlnValThrProThrSerIleuSerAlaGlnIlnThr--ProProAsnValGlnIle 112
Qy 526 GTGAGAGGACACACCTGTGTCAGACCAATCGAGCGGATGCGCCCTTCAACGATGTC 585
Db 112 u-ThrgIyTyrlArgValArgValThrProIySgIySthrgIyPrometIySgIlnIleA 132
Qy 586 GTATTGCTGAGACCTGAGTGCTGTAGACCGTCAATCTGCAACAACAGCAATGTGCTG 645
Db 132 snIleuAla-----ProAspSerSerSerValValI 142
Qy 646 TTAGCGTGTGCTGACTGTTGAGGGAGCAATCAAGCATTAAGCTTAAATTTTGA 705
Db 142 alSer-----GlyIleuMetValAlaThrIySthrgIyIlnValIleuIySAla 160
Qy 706 TCGATTAACATCAACGCTGCTCATAGAGAGAAAGAGAG-----CAAGCTTAAGTCAA 762
Db 160 spThrIleuThrSerArgProAlaGlnGlyValValThrThrIleuGlnAsnValSerProP 180

```

```

OY 763 AATCAAAACATTTGCTAGTAGTGGCGCGATGTGCATTAACCTTGAGAAAGCTGACT 822
Db      |||
OY 180 roArgVal1aArgVal1ThrAsp----- 187
OY 823 TACTAAAGCTATTCAAGAACATTTGATCGCTAACGTCACAGAACGACTACTTTG 882
Db      -----
OY 187 ----- 187
OY 883 AGTCATTTGATTGGACAGATGCAACCTTACTGTGAAAGCGAAGTCTACTTTG 942
Db      |||
OY 188 -----AlaThrGluThrThrIleThrIle----- 195
OY 943 CTGACAAAGATGGTTCGGTAACCTTCCGACCAACCTGTCCAGAAATTTTGCTAACG 1002
Db      |||
OY 196 -----SerTyrArgThrIleThrGluThrIleThrGluPheGlnValAsp 211
OY 1003 GACATGTGGCGCTTAGACCATATTAAGAAAACCAATACAAAACCAAGCAATCTGTG 1062
Db      |||
OY 211 la-----ValProAlaAsnGlyGlnThrProIleGlnArgThrIleLysProAsp 228
OY 1063 ATGTGAATATATCTGTACAGTTTACTCCCTTAACCTGTGACGATTTCAGACCGATC 1122
Db      |||
OY 228 laArgSerTyrThrIle-----ThrGlyLeuGlnProGlyThrAspTyrLys----- 243
OY 1123 TCAAGATATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAA 1182
Db      |||
OY 244 -----IleTyrLeuTyrThrLeuAsnAspAsnAlaArgSerSerPro 258
OY 1183 TACTAGCTCAAGCA----- 1196
Db      |||
OY 258 aValAlaIleAspAlaSerThrAlaIleAspAlaProSerAsnLeuArgPheLeuAlaThr 278
OY 1197 -----CAAAAGCATTTTAAACAAACAAACCA-----GGCTATATCA 1233
Db      |||
OY 278 hrProAsnSerLeuLeuValSerTyrGlnProProAlaArgAlaIleThrGlyTyrIle 288
OY 1234 TT-----TATGAACGTGACTCTCAATCGTCACTGATGACATGACATTTCCGTACAT 1290
Db      |||
OY 298 leLysTyrGluYArgLysProGlySerProPro-----ArgGluVal 311
OY 1291 TACCAATGATCA-----G 1305
Db      |||
OY 311 alProArgProArgProGlyValThrGluAlaThrIleThrGlyLeuGlnProGlyThrG 331
OY 1306 AGTTACTACTACCGTGTAAATAATCGGAAACACCTTATAGATCAATAAAATCT----- 1361
Db      |||
OY 331 lutyThrIleTyrVal-----IleAlaLeuValAsnAsnGlnLysSerGlu 347
OY 1362 -----GGTCTGAATGAAGAAATTAACAACACTGACCTGATCTCGAAGAAATATTAC 1413
Db      |||
OY 347 roLeuIleGlyArgGlyLysThrValGlnLysThrProPheValThrHisProGlyTyr 367
OY 1414 TCCTTAAAAAGG----- 1427
Db      |||
OY 367 spThrGlyAsnGlyIleGlnLeuProGlyThrSerGlyGlnProSerValGlyGln 387
OY 1427 ----- 1427
OY 387 lmeIlePheGluGluHisGlyPheArgArgThrThrProThrThrAlaThrPro 407
OY 1428 -----GAAAGCCGTATGATCCC-----TTTATGCGCACTGACTTGAACCTGTCA 1473
Db      |||
OY 407 leArgHisArgProArgProTyrProProAsnValGlyGlnGluAlaLeuSerGlnThr 427
OY 1474 CCATCAAAATACGTTGATGTC-----GATACCAACGAATGCTAAAGAGAGACCTGTAA 1530
Db      |||
OY 427 hlllSerThrAlaProPheGlnAspThrSerGlyTyrIleIleSerCysHisProVal 447
OY 1531 CAGTAGCAACAGCTAATCTTGAATTCAGAGATTATACATCTCGTGAATAGGCTTAAC 1590
Db      |||
OY 447 lYThrAspGluGluProLeuGlnPheArg-----ValProGlyThrSerThrSer 464
OY 1591 TACTCTAACCAACATCTCGATGCTTTGGATTATAGACTATACCTTAACCTGGA----- 1643

```

```

Db 464 la-----ThrLeuThrGlyLeuThr 471
OY 1644 -----AAAGTAGAGATTAATCAACATGACA 1668
Db 471 rglValaThrTyrAsnIleIleValGluAlaLeuLysAspGlnGlnArgHisLysVal 491
OY 1669 CCAACCGTATCATTAACGTTTATATGCGCAAGCGAACCGAAGAGAGAAATGCTAGTACC 1728
Db      |||
OY 491 rglGluGluValaThrValGlyAsnSerValAsn-----GluGlyLeuAsnGln----- 507
OY 1729 ATTTAGCTGTGTGCGCAGCGGCAACAGATTGATCCCATGATGACGAAAGTGTTCATC 1788
Db      |||
OY 508 -----ProThrAspArgSerCysPheAsp 516
OY 1789 ATGCTGTGGGACTCTTATGTGTCGGAGAAAGCTGGAGAAAGCCATCAAGGCTGGA 1848
Db      |||
OY 516 roTyrThrValSerHisThrValaValaGlyAspGluTyrGluArgMetSerGluSerGly 536
OY 1849 TGATGTGATGATTGTACTTGTCTGGAGAGAGCGGACGCGATCACTTGCACTTGAA 1908
Db      |||
OY 536 helYsLeuLeuCysGlnCysLeuGlyPheGlySerGlyHisPheArgCysAspSerSer 556
OY 1909 ATAGATCAACGATCAGAGACACAAAGACATCTTATAGATTGAGACACTTGACAGACA 1968
Db      |||
OY 556 rgtProCysHisAspAsnGly-----ValAsnTyrLysIleGlyLysThrAspArg 574
OY 1969 AGGATTAATCGAGAAACCTGCTCCAGTGCATCTGCACAGGACGCGGAGAGATGGA 2028
Db      |||
OY 574 lngLysGluAsnGlyGlnMetMetSerCysThrCysLeuGlyAsnGlyLysGlyGluPhe 594
OY 2029 AGTGTGAGAGGCAC 2042
Db      |||
OY 594 yscYAspProHis 598

RESULT 15
US-10-995-561-630
; Sequence 630, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 630
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-630

Alignment Scores:
Pred. No.: 1,71e-13 Length: 1315
Score: 275.50 Matches: 142
Percent Similarity: 32.88% Conservative: 77
Best Local Similarity: 21.32% Mismatches: 220
Query Match: 7.39% Indels: 227
DB: Gaps: 25

US-09-940-235-12 (1-2096) x US-10-995-561-630 (1-1315)
OY 346 TATGTGTGGAGAAAGTGGAGAAAGCGAGCCATCACTTGCACTTGAATA 405
Db      |||
OY 683 TyrValValSerValTyrAlaGlnAsnProSerGlyLysArgGlnProLeuValGlnThr 702
OY 406 GATGACAGATTCAGACACAGACATCTATAGATTGAGACACCTGAGACAAAG 465
Db      |||
OY 703 AlaValThr-----ThrIleProAlaProThrAspLeuLys 714

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:49:35 ; Search time 32.1209 Seconds

(with 953 Million cell updates/sec)

Title: US-09-940-235-12

Perfect score: 3726

Sequence: 1 cgaagaccatcatgtgtcgtt.....ccttcacgatgtcgttag 2096

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Dgapop 6.0 , Dgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
MODEL=frame+ n2p.model -DEV-xlp
-O=/cgn2.1/USPTO_spool_p/US09940235/rnatc_27012006_144218_27578/app.query.fasta_1.7708
-DB=PIR -QFMT=fastan -SUFIX=rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIG=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USERS=US09940235 @CCN_1_1_185 @runatc_27012006_144218_27578 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJOURN -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
1: pir_80:*
2: pir1:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1984	53.2	440	1 A22801	streptokinase prec
2	1951.5	52.4	415	1 BZSO	streptokinase A (E)
3	1942	52.1	440	2 S02723	streptokinase G pr
4	1779	47.7	440	2 S02724	streptokinase A pr
5	1673	44.9	414	2 JU0292	streptokinase A pr
6	1670	44.8	440	2 S04168	streptokinase A pr
7	667	17.9	197	2 S53334	streptokinase - St
8	640.5	17.2	2385	1 FNHU	fibronectin precur
9	624.5	16.8	2265	1 FNBO	fibronectin - bovi
10	620	16.6	128	2 S77671	streptokinase A (E)
11	587	15.8	2477	2 S14428	fibronectin precur
12	569	15.3	128	2 S77680	streptokinase A (E)
13	557	14.9	128	2 S77688	streptokinase A (E)
14	554.5	14.9	2481	2 A43908	fibronectin - Afri

15	552	14.8	128	2 S77679	streptokinase A (E)
16	423	11.4	128	2 S77676	streptokinase A (E)
17	419	11.2	128	2 S77673	streptokinase A (E)
18	419	11.2	128	2 S77687	streptokinase A (E)
19	417	11.2	128	2 S77682	streptokinase A (E)
20	416	11.2	128	2 S77678	streptokinase A (E)
21	408	11.0	128	2 S77677	streptokinase A (E)
22	407	10.9	128	2 S77685	streptokinase A (E)
23	401	10.8	128	2 S77683	streptokinase A (E)
24	398	10.7	128	2 S77686	streptokinase A (E)
25	397	10.7	128	2 S77684	streptokinase A (E)
26	396	10.6	128	2 S77681	streptokinase A (E)
27	390	10.5	128	2 S77675	streptokinase A (E)
28	386	10.4	128	2 S77672	streptokinase A (E)
29	226	6.1	1020	2 A29355	fibronectin - chic
30	182	4.9	190	2 S15129	fibronectin - east
31	146.5	3.9	2231	2 D71870	hypothetical prote
32	145	3.9	132	1 QOEC8	hypothetical prote
33	138	3.7	1315	2 T28679	fibrirogen-binding
34	136.5	3.7	2334	2 S32920	cell wall-associat
35	132	3.5	1385	2 D89824	hypothetical prote
36	131	3.6	1350	2 T30294	hypothetical prote
37	130.5	3.5	940	2 AD1374	invariant surface
38	129.5	3.5	13055	2 T16580	internalin protein
39	128	3.4	2256	2 AD1018	large repetitive p
40	126.5	3.4	1512	2 T14883	hypothetical prote
41	126	3.4	4688	2 P82885	hypothetical prote
42	125.5	3.4	940	2 AB1744	internalin protein
43	125	3.4	1002	2 T02468	hypothetical prote
44	125	3.4	1220	1 DJBEC3	DNA-directed DNA p
45	125	3.4	1417	2 H90670	probable invasiv

ALIGNMENTS

```
RESULT 1
A22801
Streptokinase precursor - Streptococcus "equisimilis"
C/Species: Streptococcus "equisimilis"
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: A22801
R/Malke, H.; Roe, B.; Ferretti, J.J.
Gene 34, 357-362, 1985
A/Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H
A/Reference number: A22801; MUID:85232084; PMID:2989113
A/Accession: A22801
A/Molecule type: DNA
A/Residues: 1-440 <MAL>
A/Cross-references: UNIPROT:P00779; UNIPARC:UPI000002BE73; GB:X72832; NID:9407876; PIDN
A/Experimental source: strain H46A
C/Genetics:
A/Gene: skc
C/Superfamily: streptokinase
```

Alignment Scores:

```
Pred. No.: 3.29e-126
Score: 1984.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 53.25%
DB: 1
Gaps: 0
```

US-09-940-235-12 (1-2096) x A22801 (1-440)

QY	588	ATTCCTGACCTGAGTGTGCTGACGCTCATCTGCAACAGCAGCAATTGGTT	647
DB	27	TTCCTGACCTGAGTGTGCTGACGCTCATCTGCAACAGCAGCAATTGGTT	46
QY	648	AGCGTTCGTGTAATCTGTTGAGGGGACCAATCAACAGCAATTAATTTTGAATC	707
DB	47	TTCCTGACCTGAGTGTGCTGACGCTCATCTGCAACAGCAGCAATTGGTT	66
QY	708	GATCTAACATCAACGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCAAAATCA	767

```

Db      |||
67  AspleuthinSerArgProIahHisglYgLYsThrGluInglYleuSerProlySser 86
QY      |||
768  AAACCAATTGCTACTGATAGTGCGCGGATGTCAATAACTTGAGAAAGCTGACTTA 837
Db      |||
87  LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluIulysAlaAspLeuLeu 106
QY      |||
828  AAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGACACTTCTTGAGGTC 887
Db      |||
107  LysAlaIleGluInglYleuIleAlaAsnValHisSerAsnAspAspPyrPheGluVal 126
QY      |||
888  ATTGATTTTGAAGCGATGCAACCATTAAGTATCGAAACGGCAGGCTTACTTGGCTGAC 947
Db      |||
127  IleAspPheAlaSerThrAspAlaThrIleThrAspArgAsnGlyValIleThrPheAlaAsp 146
QY      |||
948  AAAAGATGGTTCGGTAACTTGGCCGACCCAACTGTCTCAAGAAATTTTGTACCGGACAT 1007
Db      |||
147  LysAspGlySerValThrLeuProThrGlnProValGluInglYpHeLeuSerGlyHis 166
QY      |||
1008  GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTG 1067
Db      |||
167  ValArgValArgProIyArgGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY      |||
1068  GAAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTTCAGACCAAGGCTCAAA 1127
Db      |||
187  GluYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY      |||
1128  GATATTAAGCTATTTGAAAACTAGCTATCGGTGACACCATCACTTGACAAATTAATA 1187
Db      |||
207  AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGluInglYleuLeu 226
QY      |||
1188  GCTCAAGCAACAAGCAATTTTAAACAACCAACCCAGGCTATTCGATTTATGACGTGAC 1247
Db      |||
227  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyYrThrIleYrGluAspArg 246
QY      |||
1248  TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTTCACATTTGATGATGAC 1307
Db      |||
247  SerSerIleValIleThrHisAspAsnAspIlePheAspGThrIleLeuProMetAspGlnGly 266
QY      |||
1308  TTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATTAATAATCTGGCTG 1367
Db      |||
267  PheThrYrArgValLysAsnArgGluGlnAlaYrArgIleAsnLysLysSerGlyLeu 286
QY      |||
1368  AATGAAGAATATTAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1427
Db      |||
287  AsnGluGluIleAsnAsnThrAspLeuIleSerGluLysYrYrValIleuLysLysGly 306
QY      |||
1428  GAAAAGCGGTATGATCCCTTGAATGCGACACTGGAACCTGTCACCATCAATACGTT 1487
Db      |||
307  GluLysProIyArgAspProPheAspArgSerHisLysLysLeuPheThrIleLysYrVal 326
QY      |||
1488  GATGTGATACCAACGAATTGCTTAAATAAGTAGACACTTAAACAGCTACCGAACCTTAAC 1547
Db      |||
327  AspValAspThrAsnGluLeuLeuLysSerGluInglYleuLeuThrIleAspGlnAsn 346
QY      |||
1548  TTAGAATTGAGATTTATACATCTCTGTGATAGGCTTAAACTTATCAACCAATCTTC 1607
Db      |||
347  LeuAspPheArgAspLeuYrAspProArgAspLysAlaLysLeuLeuYrAsnAsnLeu 366
QY      |||
1608  GATGCTTTTGGATATATGACATATACCTTAACCTGGAATAAGTAGAGATTAATCAAGATGAC 1667
Db      |||
367  AspAlaPheGlyIleMetAspYrThrLeuThrGlyLysValGluAspAsnHisAspAsp 386
QY      |||
1668  ACCAACCGTATCATTAACGTTTATATAGGCAAGCGAACCGAAGAGAGATGCTAGCTAC 1727
Db      |||
387  ThrAsnArgIleIleThrValYrMetGlyLysArgProGluGluAsnAlaSerYr 406
QY      |||
1728  CATTTAGCT 1736
Db      |||
407  HisLeuAla 409

```

RESULT 2
BZSO

```

streptokinase (EC 3.4.-.-) - Streptococcus sp.
C:Species: Streptococcus sp.
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C:Accession: A00967
R:Jackson, K.W.; Tang, J.
Biochemistry 21, 6620-6625, 1982
A:Title: Complete amino acid sequence of streptokinase and its homology with serine proteases
A:Reference number: A00967; PMID:83127125; PMID:6760891
A:Accession: A00967
A:Molecule type: protein
A:Residues: 1-415 <JAC>
A:Cross-references: UNIPROT:P00779; UNIPARC:UPI0000172BFS
A:Note: 169-Asp and 181-Asp were also found
A:Note: this protein is not a protease, but it activates plasminogen by complexing with C:Superfamily: streptokinase
C:Keywords: hydrolase

Alignment Scores:
Pred. No.: 5,1e-124 Length: 415
Score: 1951.50 Matches: 380
Percent Similarity: 99.22% Conservative: 1
Best Local Similarity: 98.96% Mismatches: 2
Query Match: 52.38% Indels: 1
DB: 1 Gaps: 1

US-09-940-235-12 (1-2096) x BZSO (1-415)
QY      |||
588  ATTGCTGACCTGATAGTGCTGCTAGACCGTCACTGTGTCAACAACCAAGCAATGGTTGTT 647
Db      |||
1  IleAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGluInglYVal 20
QY      |||
648  AGCGTTGCTGTACTGTGTTGAGGGGAGCATCAAGACATTAGTCTTAATTTTGAATC 707
Db      |||
21  SerValAlaGlyThrValGluGluYrThrIleAsnGlnAspIleSerLeuLysPheGluIle 40
QY      |||
708  GATCTTAACATCAAGACCTGCTCATGAGGGAAGCAAGAGCAAGAGGCTTAAGTCAAAATCA 767
Db      |||
41  AspleuthinSerArgProIahHisglYgLYsThrGluInglYleuSerProlySser 60
QY      |||
768  AAACCAATTGCTACTGATAGTGCGCGGATGTCAATAACTTGAGAAAGCTGACTTA 827
Db      |||
61  LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 80
QY      |||
828  AAGGCTATTCAAGAACAAATTGATGCTTACGTCACAGTAACGACTTCTTGAGGTC 887
Db      |||
81  LysAlaIleGluInglYleuIleAlaAsnValHisSerAsnAspAspYrPheGluVal 100
QY      |||
888  ATTGATTTTGAAGCGATGCAACCATTAAGTATCGAAACGGCAAGGCTTACTTGGCTGAC 947
Db      |||
101  IleAspPheAlaSerThrAspAlaThrIleThrAspArgAsnGlyLysValYrPheAlaAsp 120
QY      |||
948  AAAGATGGTTCGGTAACTTGGCCGACCCAACTGTCCAAAGAAATTTTGTAAAGGACAT 1007
Db      |||
121  LysAspGlySerValThrLeuProThrGlnProValGluInglYpHeLeuSerGlyHis 140
QY      |||
1008  GTGCGCGTTAGACCATATTAAGAAAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1067
Db      |||
141  ValArgValArgProIyArgGlyLysProIleGlnAsnGlnAlaLysSerValAspVal 160
QY      |||
1068  GATATTAAGCTATGATTTTAACTCCCTTAAACCTGACGATGACGATTTTCAACAGGCTCAAA 1127
Db      |||
161  GluYrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
QY      |||
1128  GATATTAAGCTATTTGAAAACTAGCTATCGGTGACACCATCACTTCAAGATTAATA 1187
Db      |||
181  LeuThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGluInglYleuLeu 200
QY      |||
1188  GCTCAAGCAACAAGCAATTTTAAACAACCAACCCAGGCTATTCGATTTATGACGTGAC 1247
Db      |||
201  AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyYrThrIleYrGluAspArg 220
QY      |||
1248  TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTTCACATTTGATGATGAC 1307

```

Db 221 SerSerIleValIthrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 240
Qy 1308 TTACTTACCGGTGTAAATAATCGGGAACAAGCTTATAGATCAATAATAAATCGTGTG 1367
Db 241 PheThrTyrArgValIlysaAsnArgGlnGlnAlaTyrArgIleAsnIlyblySerGlyLeu 260
Qy 1368 AATGAAGAAATTAACAACACTGACCTGATCTTGAGAAATAT--TACGTCTTAAAAA 1424
Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerLeuGlnTyrIlyblyrValIleuIlybly 280
Qy 1425 GGGGAAAAGCCGATATATCCCTTGAATCGAGATCACTTGAACCTGTCAACATATAC 1464
Db 281 GlyGlnIlyblyProTyrAspProPheAspArgSerHisLeuIlyblyPheThrIleIlyblyr 300
Qy 1485 GTTGATGTGATACCAACGAATGTCTTAAAGTAGAGAGCTTTAACAGTAGAGAAAGT 1544
Db 301 ValAspValAspThrAsnGlnLeuLeuIlyblySerGlnIleuLeuThrAlaSerGlnArg 320
Qy 1545 AACTTAGACTTCAGAGATTATACGATCCTCGTATTAAGGCTTAACTACTTACACAAAT 1604
Db 321 AsnLeuAspPheArgAspLeuTyrAspProArgAspIlyblyAlaIlyblyLeuTyrAsnAsn 340
Qy 1605 CTCGATGCTTTGGTATATATGACATTAACCTTAACTGAAAAGTAGAGATATACGAT 1664
Db 341 LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyIlyblyValGlnAspAsnHisAsp 360
Qy 1665 GACACCAACCGTATCATTAACCGTTTATATGGGCAAGGCAACCGGAAGAGAAATGCTAC 1724
Db 361 AspThrAsnArgIleIleThrValTyrMetGlyIlyblyArgProGlnGlnIlyblyAsnAlaSer 380
Qy 1725 TACCATTTAGCT 1736
Db 381 TyrHisLeuAla 384
RESULT 3
S02723
Streptokinase G precursor - Streptococcus sp. (strain 19908)
C:Species: Streptococcus sp.
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02723
R:Walter, F.; Siegel, M.; Walke, H.
Nucleic Acids Res. 17, 1262, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.
A:Reference number: S02723; MID:89160265; PMID:2922269
A:Accession: S02723
A:Molecule type: DNA
A:Residues: 1-440 <MAL>
A:Cross-references: UNIPROT:P10519; UNIPARC:UPI000013614B; EMBL:X13400; NID:g47095; PIR
C:Genetics:
A:Gene: skg
C:Superfamily: streptokinase
F1-26/Domain: signal sequence #status predicted <SIG>
F1-27-440/Product: streptokinase #status predicted <MAP>
Alignment Scores:
Pred. No.: 2,256-123 Length: 440
Score: 1942.00 Matches: 375
Percent Similarity: 98.43% Conservative: 2
Best Local Similarity: 97.91% Mismatches: 6
Query Match: 52.12% Indels: 0
DB: 2 Gaps: 0
US-09-940-235-12 (1-2096) x S02723 (1-440)

Db 67 AspLeuThrSerArgProAlaHisGlyGlyblyblyThrGlnGlnGlnIlyblySerProIlyblySer 86
Qy 768 AAACATTTGCTACTGATAGTGGCGGATGTCACATTAACCTTGAGAAAGCTGACTTACTA 827
Db 87 LysLeuPheAlaThrAspSerGlyAlaMetProHisblyblyGlnIlyblyAlaAspLeuLeu 106
Qy 828 AAGCTATTCAAGAACATTTGATCGCTACCGTCCACAGTAACGACGACTTGTAGGCTC 887
Db 107 LysAlaIleGlnGlnGlnIleuIleAlaAsnValHisSerAsnAspAspTyrPheGlnVal 126
Qy 888 ATGATTTTGGCAAGCATGCAACCATTAATGATCGCAACCGCAAGTCTACTTGTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyIlyblyValTyrPheAlaAsp 146
Qy 948 AAAGATGTTGGTATACCTTGGCCGACCACTGTCCAAAGATTTTGTGTAAGCGACAT 1007
Db 147 LysAspGlySerValIthrLeuProIleGlnProValGlnGlnPheLeuLeuIlyblyHis 166
Qy 1008 GTGGCGCTTAGACCATATTAAGAAACCAATPACAAACCAAGCAAACTGTTGATGTG 1067
Db 167 ValArgValArgProTyrTyrGlnIlyblyProValGlnAsnGlnAlaIlyblySerValAspAl 186
Qy 1068 GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACACCAAGTCTCAA 1127
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProAlaLeuIlybly 206
Qy 1128 GATATCAAGCTATTTGAAAACATGACTATCGGTGACACCATCACATCTCAAGATTTACTA 1187
Db 207 AspThrIlyblyLeuLeuIlyblyThrLeuAlaIleGlyAspThrIleThrSerGlnIlyblyLeu 226
Qy 1188 GCTCAAGACAAACATTTTAAACAAACCAACGAGCTTATGATTTTAAGACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnIlyblyAsnHisblyblyGlyTyrThrIleTyrGlnArgAsp 246
Qy 1248 TCCCTCAATCGTCACTGACATGACATGATCTTTCGTAAGATTTTACCAATGATGATCAAG 1307
Db 247 SerSerIleValIthrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy 1308 TTACTTACCGGTGTAAATAATCGGGAACAAGCTTATAGATCAATAATAAATCGTGTG 1367
Db 267 PheThrTyrHisValIlysaAsnArgGlnGlnAlaTyrArgIleAsnIlyblySerGlyLeu 286
Qy 1368 AATGAAGAAATTAACAACACTGACCTGATCTTGAGAAATTTACGTCTTAAAAAGGG 1427
Db 287 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnIlyblyTyrValIleuIlyblyGly 306
Qy 1428 GAAAAGCGTATGATCCCTTGAATCGAGTCACTGAAACGTTCACATCAATTAAGTT 1487
Db 307 GlnIlyblyProTyrAspProPheAspArgSerHisLeuIlyblyPheThrIleIlyblyrVal 326
Qy 1488 GATGTCGATACCAACGAATGTCTTAAAGTAGAGAGCTTTAACAGTAGAGCAAGTATAC 1547
Db 327 AspValAsnThrAsnGlnLeuLeuIlyblySerGlnIleuLeuThrAlaSerGlnArgAsn 346
Qy 1548 TTAGACTTCAGAGATTATACGATCCTCGTATTAAGGCTTAACTACTTACACAAATCTC 1607
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspIlyblyAlaIlyblyLeuTyrAsnAsnLeu 366
Qy 1608 GATGCTTTGGTATTAAGACTATACCTTAACTGAAAAGTAGAGAGATTAACAGATGAC 1667
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyIlyblyValGlnAspAsnHisAspAsp 386
Qy 1668 ACCAAGCGTATCATTAACCGTTTATATGGGCAAGGCAACCGGAAGAGAAATGCTAGCTAC 1727
Db 387 ThrAsnArgIleIleThrValTyrMetGlyIlyblyArgProGlnGlnIlyblyAsnAlaSerTyr 406
Qy 1728 CATTTAGCT 1736
Db 407 HisLeuAla 409
RESULT 4
S02724

streptokinase A precursor - Streptococcus pyogenes (strain SF130/13)
C.Species: Streptococcus pyogenes
C.Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 09-Jul-2004
C.Accession: S02724
R.Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1261, 1989
A.Title: Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes tye
A.Reference number: S02724; MUID:89160264; PMID:2646590
A.Accession: S02724
A.Molecule type: DNA
A.Residues: 1-440 <MWL>
A.Cross-references: UNIPROT:P10520; UNIPARC:UPI000017021D; EMBL:X13399; NID:947435; PION
C.Genetics:
A:Gene: ska
C:Superfamily: streptokinase
F:1-26/Domains: signal sequence #status predicted <SIG>
F:27-440/Product: streptokinase #status predicted <MAT>

Alignment Scores:

Pred. No.:	2,26e-112	Length:	440
Score:	1779.00	Matches:	338
Percent Similarity:	93.73%	Conservative:	21
Best Local Similarity:	88.25%	Mismatches:	24
Query Match:	47.75%	Indels:	0
DB:	2	Gaps:	0

US-09-940-235-12 (1-2096) x S02724 (1-440)

QY	588	ATTGCTGCACCTGATGGTGTCTAGACCGTCCTCATCTGTCAACAACAGCCAAATTTGTTGTT	647
Db	27	lealIGllyrteglYrrpleuProspBprgProProlleamnsSerGlmeuValval	46
QY	648	ACGCTTGCTGTACTGTTTGAAGGGACGAAATCAGAATCAAGACTTAAGCTTTAAATTTTTGAATC	707
Db	47	sermetAlagIyleValGIuglYThrAspyslysValPheileamPheheguille	66
QY	708	GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATA	767
Db	67	AapleuthrseroinProAlahIsGlYglyThsrngluInglyLeuserrProlysser	86
QY	768	AAACCAATTTGCTACTGATAGGGGGCGATGTCACATPAACCTTGAGAAAGGACTTACTA	827
Db	87	LysProPheAlathrsaspaenglyAlamecProhlslsyleuGlulysAlaspHeuleu	106
QY	828	AAAGCTATTCAGAACCAATTGATCGCTAACGCTCACAGTAACAGACGACTTCTTGAGGTC	887
Db	107	LysAlaIlelglnLysglnleuIlleAlasnValHisserAmnAspGlyTrPhegluVal	126
QY	888	ATTGATTTTGCAGAGCGATGCAACCATTAAGTGAAGAAAGCGCAAGGCTTACTTGTGCAC	947
Db	127	IleasppheAlaseraspAlatrllleThraspargaenglylsValTrypheAlaasp	146
QY	948	AAAGATNGTTCGGTAACCTTGCGGACCCCAACCTGCCAAGATTTTCTTAAGCGGACAT	1007
Db	147	LysAspGlySerAlthrleuProThrglnProValGlnglnPheValLysglYhis	166
QY	1008	GTGCGGCTTAGACCATTAATAAAGAAAAACAATCAAAAACAAGCGCAATCTGTTGATG	1067
Db	167	ValargValArgproTryLysglulysProValGlnmenglnAlalySerValaspVal	186
QY	1068	GAATATATCTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTGACACGAGGTCCAA	1127
Db	187	GlutyTrThValGlnPheThrProleuasnpRobspaspaepheArgProglyLeuyls	206
QY	1128	GATACCTAACCTAATTGAAAAACATAGCTATGGGTGACACATCAACATCTCAAGAAATTACTA	1187
Db	207	AspThrLysleuLeuylstrlleuAlaleIyAspThrllleThrsSerglInleuLeu	226
QY	1188	GCTCAAGCAACAAGCAATTTTAAACAANAACAACGAGGCTATACGATTATANGAAGTGAC	1247
Db	227	AlaglnAlaglnserlleleuasenLysThrhIsproglYtryThmlletryGlnUargasp	246
QY	1248	TCTCGAATGCTCATCATGACAAATGACATTTTCCGTACGATTTTCAATGAGATCAAG	1307

Db 247 SetSerIleValThrHishePasnAspIlePheArgThrIleLeuProctAspGlnGlu 266

QY 1308 TTTACTTACCGTGTAAAAATCGGAAACAAGCTTATAGATCAATAAAAAATCTGGCTCG 1367

Db 267 PheThrTyHisValIysAsnArgGluGlnIleValIleAsnProLysThrGlyIle 286

QY 1368 AATGAAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAGGG 1427

Db 287 LysGluIleThrAsnAsnThrAspLeuValSerGluIleValIleValIleValIleValIle 306

QY 1428 GAAAGCCGATGATCCCTTGTATGCGAGTCACTTGAACCTTGAACCTTGAACCTTGAACCTT 1487

Db 307 GluLysProLysAspProPheAspArgSerHisLeuLysLeuPheThrIleLysIleVal 326

QY 1488 GATGTGATACCAACGAATGCTTAAAAAGTAGAGCAGCTTAAACAGCTACCGCAACCTAAC 1547

Db 327 AspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluGlnLys 346

QY 1548 TTAGACTTCAGAGATTATTAACGATCCTCGTATTAAGGCTTAAACTACTACCAAACTTC 1607

Db 347 LeuAspPheArgAspLeuIleValIleValIleValIleValIleValIleValIleValIle 366

QY 1608 GATGCTTTGGTATATATGACTATACCTTAACTGGAAGAATGAGATATATACCAATGAC 1667

Db 367 AspAlaPheAspIleMetAsnIleValIleValIleValIleValIleValIleValIleVal 386

QY 1668 ACCAACCCGATCAATCAACCGTTATATGAGGCAAGCAACCGCAAGAGAGAGATGCTAGCTAC 1727

Db 387 AsnAsnAspValValIleValIleValIleValIleValIleValIleValIleValIleVal 406

QY 1728 CATTTAGCT 1736

Db 407 HisLeuVal 409

RESULT 5

JU0292

Streptokinase - Streptococcus pyogenes

C:Species: Streptococcus pyogenes

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C:Accession: A43867, JU0292

R:Okumura, H.; Todome, Y.; Suzuki, H.; Mitsu, M.; Kotani, N.; Horiuchi, K.; Shikama, N.;

I:Infect. Immun. 60, 278-283, 1992

A:Title: Immunochemical studies and complete amino acid sequence of the streptokinase fr

A:Reference number: A43867, PMID:92104686; PMID:1370275

A:Accession: A43867

A:Molecule type: protein

A:Residues: 1-414 <OHK>

A:Cross-references: UNIPROT:Q57391; UNIPARC:UP10000175C82

A:Experimental source: M type 12 strain A374

A:Note: sequence extracted from NCBI backbone (NCBI:74592)

C:Superfamily: streptokinase

Alignment Scores:

Pred. No.: 3,19e-105 Length: 414

Score: 1673.00 Matches: 324

Percent Similarity: 90.60% Conservative: 23

Best Local Similarity: 84.60% Mismatches: 36

Query Match: 44.90% Indels: 0

DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x JU0292 (1-414)

QY 588 ATTTGTCGACCTGAGTGGCTGCTAGACCGTCCATCTGCAACAACGCAATGCTGTT 647

Db 1 IleAlaGlyProGluIleValLeuAspArgProSerValaAsnSerGlnLeuValVal 20

QY 648 AGCGTTGCTGCTACTGTGAGGGGACGAATCAACATTAAGCTTAAATTTTGGAAATC 707

Db 21 SerValAlaGlyThrValGluGlyThrAsnGlnGluIleSerLeuLysPheHeGluIle 40

QY 708 GATTCATCAATCAACGCTGCTCATGAGAAAGACAGACAGGCTTAAGTCCAAAATCA 767

```

Db      41  AspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGlyLeuSerProLysSer 60
Qy      768  AAACCATTTGCTACTGATAGTGGGCGCATGTCACATAACTTGAGAAAGTGACTTACTA 827
Db      61  LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeu 80
Qy      828  AAGGCTATTTCAGAAACATTTGATCGCTAACCTGCACAGTAAAGACGACTTGTGAGTTC 887
Db      81  LysAlaAlaGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAsnArgLysTyrPheGluVal 100
Qy      888  ATTGATTTTGCAGCCGATGCAACCATTTACTGATGGAACCGGCAAGGCTTACTTGGCTGAC 947
Db      101  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 120
Qy      948  AAAGATGATCGGTGTAACCTTGCGGACCGAACCTGTCAGAAATTTTGTGTAAGGCGACAT 1007
Db      121  ArgAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 140
Qy      1008  GTGGCGGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTG 1067
Db      141  ValArgValArgProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 160
Qy      1068  GAATATTAAGTATGACTTTACTCCCTTAAACCTGATGACGATTTTCAGACGAGTTCGAAA 1127
Db      161  AsnTyrGluValSerPheValSerGluThrGlyAsnLeuAspPheThrProSerLeuLys 180
Qy      1128  GATTAAGTATGATTAAGAAACCTAGCTATGCTGATGACCATCATCATCTGCAAGATTTACTA 1187
Db      181  GluArgTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerSerGlnGluLeuAla 200
Qy      1188  GCTCAAGCACAAGACATTTTAAACCAAAACCAACCGCTATGACTTATGACTTATGACGTAC 1247
Db      201  AlaIleAlaGlnPheIleLeuSerLysGluHisProAspTyrIleIleThrLysArgAsp 220
Qy      1248  TCTCAATGCTCATCTATGACCAATGACATTTTCCGATCGAATTTTACCAATGATGACAAAG 1307
Db      221  SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetArgGlnGlu 240
Qy      1308  TTTACTTACCGTGTAAAAATCGGGAACAGGCTTATGATCATATAAAAAATCGGTCTG 1367
Db      241  PheThrTyrHisIleLysAspArgGluGlnAlaTyrLysAlaAsnSerLysThrGlyIle 260
Qy      1368  AATGGAATTAATAACAACATGACCTGATCTGTGAGAAATATTGATGCTTAAAAAAAGG 1427
Db      261  ValGluLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValIleLysLysGly 280
Qy      1428  GAAACCCGATATGATCCCTTGTATCGCATGCTTGAACCTGTTACCATCAATATGATT 1487
Db      281  GluGluProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
Qy      1488  GATGTGATATCCACGAAATTTGCTAAAAAGTAGCAGCTGTTAACAGCTTAGCGAAAGTTAC 1547
Db      301  AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGluArgAsn 320
Qy      1548  TTAGATTTAGAGATTATATGATCCCTCGATGATAGGCTAAACATCTTCAACAACATCTC 1607
Db      321  LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
Qy      1608  GATGCTTTGTATATGACTATATCTTAACTGAAAGAGTAAAGTAAATGATCAGATGAC 1667
Db      341  AspAlaPheGlyLysMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsnAsp 360
Qy      1668  ACCCAACCGTATCATAAACGTTATATGAGGCAAGCAGCCGAAAGAGAAATGCTAGCTAC 1727
Db      361  ThrAsnArgGlleIleThrValTyrMetGlyLysArgProGluGluAsnAlaSerTyr 380
Qy      1728  CATTTAGCT 1736
Db      381  HisLeuAla 383

```

```

C:Species: Streptococcus pyogenes
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: S04168
R:Huang, T.T.; Malke, H.; Ferretti, J.J.
Mol. Microbiol. 3, 197-205, 1989
A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escheri-
A:Reference number: S04168; MUID:89343623; PMID:2668686
A:Accession: S04168
A:Molecule type: DNA
A:Residues: 1-440 <HUA>
A:Cross-references: UNIPROT:O57391; UNIPARC:UPI00000BD04A; EMBL:X51517; NID:947437; PID:
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-440/Product: streptokinase A #status predicted <MAT>

Alignment Scores:
Pred. No.: 5,13e-105 Length: 440
Score: 1670.00 Matches: 324
Percent Similarity: 90.34% Conservative: 22
Best Local Similarity: 84.60% Mismatches: 37
Query Match: 44.82% Indels: 0
Gaps: 0

US-09-940-235-12 (1-2096) x S04168 (1-440)

Qy      588  ATTGCTGACCTGATAGTGGCTGCTAGACCGTTCATCTGTCACAAACGCAATGTGTTGT 647
Db      27  IleAlaGlyTyrGluTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
Qy      648  AGCGTGTGCTGATCTGTTGAGGCGCAATCAAGCATTTAGCTTAAATTTTGAATC 707
Db      47  SerValaAlaGlyThrValGlnGlyThrAsnGlnGluIleSerLeuLysPhePheGluIle 66
Qy      708  GATTAACATCAGACGCTGCTCATGAGAGAAAGACAGAGCAAGGCTTAAGTCAAAATCA 767
Db      67  AspLeuThrSerArgProAlaHisIleGlyLysThrGlnGlnGlyLeuSerProLysSer 86
Qy      768  AAACCATTTGCTACTGATAGTGGCGCATGTCACATAACTTGAGAAAGTGACTTACTA 827
Db      87  LysProPheAlaThrAspLysGlyAlaMetSerHisLysLeuGluLysAlaAspLeu 106
Qy      828  AAGGCTATTCAAGAACATTTGATGCTTAACCTCCACAGTAAAGACGACTTGTGAGTTC 887
Db      107  LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
Qy      888  ATTGATTTTGCAGCGATGCAACCATTTACTGATGGAACCGGCAAGGCTTACTTGGCTGAC 947
Db      127  IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
Qy      948  AAAGATGATCGGTGTAACCTTGCGGACCGAACCTGTCAGAAATTTTGTGTAAGGCGACAT 1007
Db      147  LysAspAspSerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
Qy      1008  GTGGCGGTTAGACCATATTAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTG 1067
Db      167  ValArgValLysProTyrGlnProLysAlaValHisAsnSerAlaGluArgValAsnVal 186
Qy      1068  GAATTAAGTATGACTTTACTCCCTTAAACCTGATGACGATTTTACGACGAGTTCGAAA 1127
Db      187  AsnTyrGluValSerPheValSerGluThrGlyAspLeuAspPheThrProLeuLeuArg 206
Qy      1128  GATTAAGTATGATTAAGAAACCTAGCTATGCTGATGACCATCATCATCTGCAAGATTTACTA 1187
Db      207  AsnGlnTyrHisLeuThrThrLeuAlaValGlyAspSerLeuSerSerGlnGluLeuAla 226
Qy      1188  GCTCAAGCACAAGACATTTTAAACCAAAACCAACCGCTATGACTTATGATGACGTGAC 1247
Db      227  AlaIleAlaGlnPheIleLeuSerLysLysHisProAspTyrIleIleThrLysArgAsp 246
Qy      1248  TCTCAATGCTGATCATGACATGACATTTTCCGATCGAATTTTACCAATGATGACAAAG 1307

```

RESULT 6
 S04168
 streptokinase A precursor - Streptococcus pyogenes (strain NZ131)

```

Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
OY      1308 TTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCATATAAATAATCGGTCTG 1367
DB      267 PheThrYrHisIleIleLysAspArgGlnGlnIleTyrLysAlaAsnSerIleStrnGlyIle 286
OY      1368 AATGAGAATAAACAACAACATGACCTGATCTCTGAGAAAATATTACGTCTTAAAAAAGG 1427
DB      287 GlnGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysLysGly 306
OY      1428 GAAAAGCCGTATGATCCCTTTGATGCGACGACTTGAACCTGTTACCATCAAAATACGTT 1487
DB      307 GlnLysProIleYrAspProPheAspSerHisLeuLysLeuPheThrIleAsnTyrVal 326
OY      1488 GATGCTGATACCAACGAATTGCTTAAAGAGACGACTCTTAACAGCTAGCGAAGCTAAC 1547
DB      327 AspValAsnThrAsnLysLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsn 346
OY      1548 TTGACCTTCAGAGATTTATACGATCTCTGATTAAGGCTAAAGCTTACCTTACCAAACTTC 1607
DB      347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
OY      1608 GATGCTTTGGTATTATGACATTAACCTTAACTGGAAGAGTAGAGATTAATCAAGATGAC 1667
DB      367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 386
OY      1668 ACCAACCGTATCATACCGTTTATATGCGAAGCGAAGCGAAGAGAGAAATGCTAGTAC 1727
DB      387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAsnAlaSerTyr 406
OY      1728 CATTTAGCT 1736
DB      407 HisLeuAla 409

```

RESULT 7

S53334
 streptokinase - Streptococcus sp.
 C/Species: Streptococcus sp.
 C/Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 C/Accession: S53334
 R/Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
 Biochem. J. 304, 235-241, 1994
 A/Title: Function of streptokinase fragments in plasminogen activation.
 A/Reference number: S53334; MUID:95091634; PMID:7998939
 A/Accession: S53334
 A/Status: Preliminary
 A/Molecule type: protein
 A/Residues: 1-197 <SH1>
 A/Cross-references: UNIPROT:Q7M115; UNIPARC:UPI000017AC2C

Alignment Scores:

Pred. No.:	2.25e-37	Length:	197
Score:	667.00	Matches:	167
Percent Similarity:	44.71%	Conservative:	2
Best Local Similarity:	44.18%	Mismatches:	1
Query Match:	17.90%	Indels:	208
DB:	2	Gaps:	9

US-09-940-235-12 (1-2096) x S53334 (1-197)

```

OY      588 ATTGCTGACCTGATGCGTGTGCTAGACCGTCCATCTGTCAACAACAGCAATTGTTGTT 647
DB      1 IleAlaGlyProGlnIlePheLeuAspArgProSerValAsn----- 14
OY      648 AGCGTGTGCTGATCTGTTGAGGGGACGAATCAAGATTAAGTCTTAATTTTGAATC 707
DB      14 ----- 14
OY      708 GATTTAAATCAACGACCTGCTCTATGAGAGAAAGACAGACAGCGTTAAAGTCCAAATCA 767
DB      15 -----Ser 15
OY      768 AAACATTTGCTACTGATGATGCGCGGATGTCACATAACTTGAGAAAGCTGACTTAC 827

```

```

Db      16 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeuLeu 35
OY      828 AAGCTATTTCAGAAACAATTGATGCTTAAGCTGCACAGTAAACGACACTTGTGAGGTC 887
DB      36 LysAlaIleGlnIleThrGlnIleLeu----- 43
OY      888 ATGATTTTTCAGAGCATGCAACCATTAAGTATGATGAAACGCAAGGTCTACTTGTGAC 947
DB      44 -----AsnGlnLysValTyrPheAlaAsp 51
OY      948 AAGATGTTTCGTTAATCCTTCCGACCCCAACCTGTCCAGAAATTTGCTAAGCGACAT 1007
DB      52 LysAspIleSerValThr----- 57
OY      1008 GTGCGCGTTAGACCATTAATAAGAAAACCAATACAAACCAAGGAATCTGTGATGTG 1067
DB      58 -----GlnLysProIleGlnAsnGlnAlaLysSerValAspVal 70
OY      1068 GAATATATCTGTAGAGTTTACTCCCTTAAACCTGATGACGATTCAGACGAGTCTGAAA 1127
DB      71 Glu----- 71
OY      1128 GATATCAAGCTATTGAAACATAGCTATGCTGACACATCATCTCAAGAAATTACTA 1187
DB      71 ----- 71
OY      1188 GCTTCACGACAAAGCATTTTAAACAAACCAACCCAGCGCTATAGATTTTATGAACGTGAC 1247
DB      72 -----AsnHisProGlnLysThrIleTyrGlnLysArgAsp 82
OY      1248 TCCCTCAATGCTCATCATGACATGACATTTTCGCTGACATTTTACCAAGATCAAGAG 1307
DB      83 SerSerIleVal-----ThrIleLeuProMetAspGlnGlu 94
OY      1308 TTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAATAAATAATCGGTCTG 1367
DB      95 PheThrTyr-----SerGlyLeu 100
OY      1368 AATGAGAATAAACAACAACACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAAGG 1427
DB      101 AsnGlnGlnIleAsnAsnThrAspLeuIleSer----- 111
OY      1428 GAAAAGCCGTATGATCCCTTGAATCGAGTCACCTGGAACCTGTTACACATCAATAGTT 1487
DB      112 -----TyrVal 113
OY      1488 GATGCGATACCAACGAATTTGCTAAAGAGAGAGAGCTTTAACAGCTAGGAGACGTAAAC 1547
DB      114 AspValAspThrAsnGlnLysLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsn 133
OY      1548 TTGACCTTCAGAGATTTATACGATCTCTGATTAAGGCTTAACTCTTACCAACATCTGC 1607
DB      134 -----AspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 149
OY      1608 GATGCTTTGGTATTATGACATTAACCTTAACTGGAAGAGTAGAGATTAATCAAGATGAC 1667
DB      150 AspAlaPheGlyIleMet----- 155
OY      1668 ACCAACCGTATCATACCGTTTATATGCGAAGCGAAGCGAAGAGAGAAATGCT 1721
DB      156 -----IleIleThrValTyrMetGlyLysArgProGlnGlnLysAsnAla 170

```

RESULT 8

FNHU

fibronectin precursor [validated] - human
 N/Alternate names: fibronectin splice form ED-A

C/Species: Homo sapiens (man)
 C/Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2004
 C/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
 R/Dean, D.C.; Bowlus, C.L.; Bourgois, S.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
 A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A:Reference number: A26460; MUID:87175578; PMID:3031656
 A:Accession: A26460
 A:Molecule type: DNA
 A:Residues: 1-49 <DEA>
 A:Cross-references: UNIPROT:P02751; UNIPROT:Q14327; UNIPARC:UPI0000016A926; GB:M15801; NT
 R:Oldberg, A.; Ruoslahti, E.
 J. Biol. Chem. 261, 2113-2116, 1986
 A:Title: Evolution of the fibronectin gene.
 A:Reference number: A26284; MUID:86111901; PMID:3003095
 A:Accession: A26284
 A:Molecule type: DNA
 A:Residues: 1447-1540 <OLD>
 A:Cross-references: UNIPARC:UPI0000112B37; GB:M12549; NID:g182668
 A>Note: The authors translated the codon TTC for residue 1494 as Gpu
 R:Paolletta, G.; Henchcliffe, C.; Sebastiao, G.; Baralle, F.E.
 Nucleic Acids Res. 16, 3545-3557, 1988
 A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
 A:Reference number: S00846; MUID:88233940; PMID:3375063
 A:Accession: S03917
 A:Molecule type: DNA
 A:Residues: 1594-1767, 'V', 1769-1783 <PAO>
 A:Cross-references: UNIPARC:UPI000017432B; EMBL:X07718; NID:g31402
 A>Note: The authors translated the codon AAC for residue 1631 as Asp
 R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 FEBS Lett. 207, 287-291, 1986
 A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene;
 A:Reference number: A24854; MUID:87030929; PMID:3770201
 A:Accession: A24854
 A:Molecule type: DNA
 A:Residues: 1992-2147 <VIB>
 A:Cross-references: UNIPARC:UPI000017432C; GB:X04530; NID:g31436
 R:Gutman, A.; Yamada, K.M.; Kornblith, A.
 FEBS Lett. 207, 145-148, 1986
 A:Title: Human fibronectin is synthesized as a pre-propolypeptide.
 A:Reference number: A24476; MUID:87030890; PMID:3770189
 A:Accession: A24476
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-14, 'O', 16-38 <GUT>
 A:Cross-references: UNIPARC:UPI000017432D
 R:Kornblith, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A:Title: Primary structure of human fibronectin: differential splicing may generate at l
 A:Reference number: A91008; MUID:85284965; PMID:2992939
 A:Accession: A91008
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 32-1344, 1346-2080, 2112-2386 <KOR>
 A:Cross-references: UNIPARC:UPI000017432E; UNIPARC:UPI000017432F; GB:X02761
 R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
 A:Reference number: A93529; MUID:84272258; PMID:6462519
 A:Accession: A93529
 A:Molecule type: mRNA
 A:Residues: 973-2080, 2112-2386 <KOR>
 A:Cross-references: UNIPARC:UPI000017432F; UNIPARC:UPI0000174330; GB:X00739
 R:Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
 A:Reference number: A21011; MUID:83290929; PMID:6688418
 A:Accession: A21011
 A:Molecule type: mRNA
 A:Residues: 1434-1537 <OL2>
 A:Cross-references: UNIPARC:UPI0000174331; GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:
 R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2698-2704, 1985
 A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with re
 A:Reference number: A90495; MUID:85280409; PMID:2992573
 A:Accession: A90495
 A:Molecule type: mRNA
 A:Residues: 1594-2386 <BER>
 A:Cross-references: UNIPARC:UPI000016A928; GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:

R:Umezawa, K.; Kornblith, A.R.; Baralle, F.E.
 FEBS Lett. 186, 31-34, 1985
 A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
 A:Reference number: A22245; MUID:85231203; PMID:2985004
 A:Accession: A22245
 A:Molecule type: mRNA
 A:Residues: 1948-2067 <UME>
 A:Cross-references: UNIPARC:UPI00000046A; GB:M27589; NID:g182705; PIDN:AAA52465.1; PID
 A:Accession: B22245
 A:Molecule type: mRNA
 A:Residues: 1975-1991, 2017-2039 <UM2>
 A:Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590
 R:Seikiuchi, K.; Kloe, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
 Biochemistry 25, 4936-4941, 1986
 A:Title: Human liver fibronectin complementary DNAs: identification of two different me
 A:Reference number: I52394; MUID:87026578; PMID:3021206
 A:Accession: I52394
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
 A:Cross-references: UNIPARC:UPI000006804C; GB:M4060; NID:g182701; PIDN:AAA52464.1; PID
 R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
 A:Reference number: A21165; MUID:83221567; PMID:6304699
 A:Accession: A21165
 A:Molecule type: mRNA
 A:Residues: 2291-2386 <KOR>
 A:Cross-references: UNIPARC:UPI0000174334; GB:X00799; NID:g182681; PIDN:AAA52460.1; PID
 R:Garcia-Pardo, A.; Pearlsstein, E.; Frangione, B.
 J. Biol. Chem. 258, 12670-12674, 1983
 A:Title: Primary structure of human plasma fibronectin
 A:Reference number: A92398; MUID:84032463; PMID:6630202
 A:Accession: A92398
 A:Molecule type: protein
 A:Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
 A:Cross-references: UNIPARC:UPI0000174335
 R:Garcia-Pardo, A.; Gold, L.I.
 Arch. Biochem. Biophys. 304, 181-188, 1993
 A:Title: Further characterization of the binding of fibronectin to gelatin reveals the
 A:Reference number: S34791; MUID:93312001; PMID:8333285
 A:Accession: S34791
 A:Molecule type: protein
 A:Residues: 291-300, 551-560 <GAR2>
 A:Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337
 R:Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
 Thromb. Res. 43, 469-477, 1986
 A:Title: Two plasma fibronectin fragments with different gelatin-binding properties.
 A:Reference number: A60904; MUID:87019725; PMID:3532418
 A:Accession: A60904
 A:Molecule type: protein
 A:Residues: 293-301 <GR1>
 A:Cross-references: UNIPARC:UPI0000174338
 R:Calaycay, J.; Pande, H.; Lee, T.; Borel, L.; Stirl, A.; Shively, J.E.; Zardi, L.
 J. Biol. Chem. 260, 12136-12141, 1985
 A:Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p
 A:Reference number: A23901; MUID:86008277; PMID:3900070
 A:Accession: A23901
 A:Molecule type: protein
 A:Residues: 616-677, 'Q', 679-703, 'PT', <CAN>
 A:Cross-references: UNIPARC:UPI0000174339
 R:Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
 J. Biol. Chem. 257, 9593-9597, 1982
 A:Title: The cell attachment domain of fibronectin. Determination of the primary structu
 A:Reference number: A92386; MUID:82265604; PMID:7050098
 A:Accession: A92386
 A:Molecule type: protein
 A:Residues: 1441-1548 <PIB>
 A:Cross-references: UNIPARC:UPI0000141CD5
 A>Note: residues 1524-1527 are responsible for the cell-binding activity
 R:Garcia-Pardo, A.; Roa-Sagdo, A.; Frangione, B.
 Biochem. J. 241, 923-928, 1987
 A:Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom

A:Reference number: A32517; MUID:87241275; PMID:3599230
A:Accession: A32517

A:Molecule type: protein
A:Residues: 1589-1630,'T',1722-2058 <GAR3>
A:Cross-references: UNIPARC:UPI000017433A
R:Tresselt, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A>Title: Human plasma fibronectin. Demonstration of structural differences between the A
A:Reference number: S14357; MUID:91190085; PMID:2012601
A:Accession: S14357

A:Molecule type: protein
A:Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
A:Cross-references: UNIPARC:UPI000017433B
R.Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A>Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A:Reference number: A23891; MUID:85261459; PMID:4019516
A:Accession: A23891

A:Molecule type: protein
A:Residues: 2071-2080,2112-2356 <GAR4>
A:Cross-references: UNIPARC:UPI000017433C, UNIPARC:UPI000017433D
C:Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C:Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
action, and transformation.

C:Genetics:
A:Gene: GDB:FNI
A:Cross-references: GDB:119135; OMIM:135600
A:Map position: 2q34-2q34
A:Intron: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C:Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology;
C:Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-31/Domain: propeptide #status predicted <PRO>
F:32-2386/Product: fibronectin #status experimental <MAT>
F:552-87/Domain: fibrinogen and heparin binding <FRH>
F:552-87/Domain: fibronectin type I repeat homology <IF1>
F:597-135/Domain: fibronectin type I repeat homology <IF2>
F:141-179/Domain: fibronectin type I repeat homology <IF3>
F:186-225/Domain: fibronectin type I repeat homology <IF4>
F:231-270/Domain: fibronectin type I repeat homology <IF5>
F:308-608/Domain: collagen binding <CBR>
F:308-342/Domain: fibronectin type I repeat homology <IF6>
F:360-401/Domain: fibronectin type II repeat homology <2FI1>
F:420-461/Domain: fibronectin type II repeat homology <2FI2>
F:470-508/Domain: fibronectin type I repeat homology <IF7>
F:518-555/Domain: fibronectin type I repeat homology <IF8>
F:561-599/Domain: fibronectin type I repeat homology <IF9>

Alignment Scores:

Pred. No.:	Length:	
Score:	640.50	Matches: 2386
Percent Similarity:	36.7%	Conservative: 208
Best local Similarity:	27.62%	Mismatches: 153
Query Match:	17.19%	Indels: 324
DB:	I	Gaps: 28

US-09-940-235-12 (1-20) x FNUU (1-2386)

OY GACCAATCATGGTTGTTCCTAGCCGCAGCGCACAGCTTTTGACAGACAGAGTCGCTTACAGT--- 61

Dd ||| ::: ||| ||||| ||| ::: ::: ||| ||| |||
 122 AspySthrCyrlleqlaglyarglgylrgllleserCYsthrtllelaasnabgCys 141

OY -----TCGCTGCCGTATCGAGATTCACTTTCGCTTAACAAGAATAAGAACACC 106

Dd 142 HisgluglgyIghlnserlyr-LyslileglyAsphthrtp-----ArgArgr 157

OY 107 CCGCAGCGCTTAGCCGGGTCTCAAACAAGAGACACATCATGCGCACCCTGGCCAGGA 166

Dd 157 ohlegluThr----- 160

OY 167 CCCAAGCGCGCCGGAATCTCGATCCCGCAAAATTAAACAGACTCACTATAGGAGACCA 226

Dd 160 ----- 160

[illegible]

QY 1159 GTGACACATCATCTCAAGATTACTAGCTCAAGCAAAAGCATTTAAACAAAACC 1218
||| ||| |||
Db 407 -----LeuValGlnThrGlnGly----- 412
QY 1219 ACCGAGGCTATACGATTTATGACAGCTCCGATCGCTACTGACATGACATTT 1278
||| ||| |||
Db 413 -----GlyAsnSerAsnGlnValAlaLeuGlyHis----- 421
QY 1279 TCCGTCAGATTTTACCAATGATCAAGAGTTTACTTACCGCTGTTAAATCGGAAACAAG 1338
||| ||| |||
Db 422 -----PheProheLeuValAsnAsnHis----- 429
QY 1339 CTATAGATCAATAAATAATGCTGTGAATGAAGAAATTAACAACACTGACCTGATCT 1398
||| ||| |||
Db 430 -----AsnTyrThrAspCysThrS----- 436
QY 1399 CTGGAATATATAGCTCTTAAATAAGGAAAAAGCCGATATGATCCCTTGATCGCAATC 1458
||| ||| |||
Db 436 ergLugly-----ArgArgAspA 442
QY 1459 ACTTGAAA---CTGTTACCATCAATATACGTTGATGTCATACCAAC-----G 1503
||| ||| |||
Db 442 smelLytTPCyseLytThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysP 462
QY 1504 AATGCTAAAAAGTAGAGAGCTCTTAACAGCTAGCGAAGCTAATTAGACTTCAGAGATT 1563
||| ||| |||
Db 462 romelALAlaHlsGlnGlnLysCysThrThrAsnGln----- 474
QY 1564 TATACAGATTCCTGATTAAGGCTAACTAATCTACCAACAATCTCGATGCTTTGGTATTA 1623
||| ||| |||
Db 475 -----GlyValM 477
QY 1624 TGACATATACCTTATCTGGAATAAGTAGAGATTAATCAGATGACACCAACCGATATATA 1683
||| ||| |||
Db 477 et---TyrArgLLeLysArgGlnTPrAspLysGlnHlsAspMetGlyHlsMetMetArgC 496
QY 1684 CCGTTTATATAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1743
||| ||| |||
Db 496 ysthrCysValGly---AsnGlyArgGlyLysThrCysTyrAlaTyrSer----- 512
QY 1744 GCCAGGCGCAACAGATTGTATCCATAGCTAGCAAGAGTTTGTATCATGCTGCTGGAGATT 1803
||| ||| |||
Db 513 -----GlnLeuArgAspGlnCysLys-----ValAspAspLLeT 524
QY 1804 CCTATGTGTGTGGAGAAACCTGAGAGAAACCTTACCAAGGCTGATGATGATGATGATTA 1863
||| ||| |||
Db 524 hrTyrAsnValAsnAspThrPheHlsLysArgHlsGlnLysHlsMetLeuAsnGlySer 544
QY 1864 CTTCGCTGGATC 1923
||| ||| |||
Db 544 hrCysPheGlnGlnLysArgGlyArgTPrLysCysAspProValAspGlnCysGlnAspS 564
QY 1924 AGGACACAGAGACATCTATAGATTTGAGAGACATCTGAGAGAGAGAGAGATTAATCGAGAA 1983
||| ||| |||
Db 564 ergLutThrGlyThrPheTyrGlnLysArgSerThrPheLysTyrValHls---GlyV 583
QY 1984 ACCTGCTCAGTGCATCTTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2043
||| ||| |||
Db 583 alArgTyrGlnCysTyrCysTyrGlyArgGlyLLeGlyGlnTPrHlsCysGln----- 600
QY 2044 CCTGTGACAGACCATCGAGCGGATCGGCGCC 2078
||| ||| |||
Db 601 --ProLeuGlnThrTyrProSerSerGlyPro 611

RESULT 9

FNBO

fibronectin - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 09-Jul-2004

C/Accession: A26452; M23292

R/Skorenskaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.B.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A/Title: Complete primary structure of bovine plasma fibronectin.

A/Reference number: A26452; MUID:87054047; PMID:3780752
A/Accession: A26452
A/Molecule type: protein
A/Residues: 1-2265 <SKO>
A/Cross-references: UNIPROT:P07589; UNIPARC:UPI000012A7BE
R/Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A/Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A/Reference number: A21165; MUID:83221567; PMID:6304699
A/Accession: B21165
A/Molecule type: mRNA
A/Residues: 2170-2265 <KOR>
A/Cross-references: UNIPARC:UPI000016C30C; GB:K00800; NID:G163055; PIDN:AAA30521.2; PID
R/Petersen, T.B.; Skorenskaard, H.C.; Skorenskaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sott
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A/Title: Partial primary structure of bovine plasma fibronectin: three types of interna
A/Reference number: A23292; MUID:83117805; PMID:6218503
A/Accession: A23292
A/Molecule type: protein
A/Residues: 1-16, 'C', 18-20, 'S', 22-432; 447-463; 1367-1517; 1567-1673; 2062-2176, 'N', 2178-22
A/Cross-references: UNIPARC:UPI000017433B; UNIPARC:UPI000017433F; UNIPARC:UPI0000174340
C/Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.
C/Comment: The plasma fibronectin molecule consists of two chains, which are connected
C/Comment: Fibronectins bind cell surfaces and various compounds including collagen, fil
aling, and maintenance of cell shape.
C/Comment: Plasma fibronectin is synthesized by hepatocytes.
C/Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C/Keywords: acute phase; alternative splicing; collagen binding; duplication; extracell
F/21-241/Domain: fibrin and heparin binding <FBH>
F/21-56/Domain: fibronectin type I repeat homology <1F1>
F/110-148/Domain: fibronectin type I repeat homology <1F2>
F/155-194/Domain: fibronectin type I repeat homology <1F3>
F/200-239/Domain: fibronectin type I repeat homology <1F4>
F/217-277/Domain: fibronectin type I repeat homology <1F5>
F/277-311/Domain: fibronectin type I repeat homology <2F1>
F/329-370/Domain: fibronectin type II repeat homology <2F2>
F/389-430/Domain: fibronectin type II repeat homology <2F3>
F/439-477/Domain: fibronectin type I repeat homology <1F6>
F/487-524/Domain: fibronectin type I repeat homology <1F7>
F/530-568/Domain: fibronectin type I repeat homology <1F8>
F/578-661/Domain: fibronectin type I repeat homology <1F9>
F/688-770/Domain: fibronectin type III repeat homology <FN3A>
F/779-860/Domain: fibronectin type III repeat homology <FN3B>
F/875-957/Domain: fibronectin type III repeat homology <FN3C>
F/965-1046/Domain: fibronectin type III repeat homology <FN3D>
F/1055-1134/Domain: fibronectin type III repeat homology <FN3E>
F/1142-1227/Domain: fibronectin type III repeat homology <FN3F>
F/1235-1318/Domain: fibronectin type III repeat homology <FN3G>
F/1326-1404/Domain: fibronectin type III repeat homology <FN3H>
F/1410-1517/Domain: fibronectin type III repeat homology <FN3I>
F/1416-1502/Domain: cell attachment <CAD>
F/1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F/1493-1495/Region: cell attachment (R-G-D) motif
F/1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F/1600-1870/Domain: heparin binding <HB2>
F/1600-1882/Domain: fibronectin type III repeat homology <FN3L>
F/1692-1773/Domain: fibronectin type III repeat homology <FN3M>
F/1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F/1970-1972/Region: cell attachment (R-G-D) motif
F/1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F/1985-2216/Domain: fibrin binding <FB2>
F/2085-2124/Domain: fibronectin type I repeat homology <1F10>
F/2130-2167/Domain: fibronectin type I repeat homology <1F11>
F/2174-2209/Domain: fibronectin type I repeat homology <1F12>
F/1/Modified site: pyroliidone carboxylic acid (Gln) #status experimental
F/3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experiment
F/21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,271-304,302-3
7,2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted
F/1205,497,511,846,976,1213,1987/Binding site: carbohydrate (Asn) (covalent) #status exp
F/1205,1692/Binding site: carbohydrate (Asn) (covalent) #status absent
F/1943,1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F/2246/Disulfide bonds: interchain (to 2250) #status predicted
F/2250/Disulfide bonds: interchain (to 2246) #status predicted

Db 492 leThrTyAsnValAsnSerThrPheHisLeuArgHisGluGluGlnHisMetLeuAsn 512
 QY 1861 GTACTTGCCTGGAGAGAGCAGCAGCATCTTCACTTGTAGAAATAGATGCAAG 1920
 Db 512 ySThrCysPheHisGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 532
 QY 1921 ATCAGACCAAGACGACATCTTATGAAATTTGAGACACCTTGAGCAAGAAAGATTAATCAG 1980
 Db 532 spSerGluThrArgThrPheThrGlnHisLeuAspSerTrp--GluLysGlyLeuGln 551
 QY 1981 GAACCTGCTCCAGGACATCTGACAGCAGCAGCAGGAGAGAGTGGAACTGTGAGAGCC 2040
 Db 551 lValAlaGlyTrpGlnCysTyGlyArgGlyLysGlyLysGlyLysGlyLysGlyLysGln---- 569
 QY 2041 AACCTCTGTGACAGACATGACAGCAGGAGATCTGAGCC 2078
 Db 570 -----ProLeuGlnThrTyProAspThrSerGlyPro 580
 RESULT 10
 S77671
 Streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
 C:Species: Streptococcus pyogenes
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
 C:Accession: S77671; S77672
 R:Kaput, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whitlam, T.S.; Sawyer, S.A.; Musser, M.I. Microbiol. 16, 509-519, 1995
 A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus A:Reference number: S77671; MUID:96037795; PMID:756511
 A:Accession: S77671
 A:Status: translation not shown
 A:Residues: 1-128 <KAB>
 A:Molecule type: DNA
 A:Cross-references: UNIPARC:UPI0000000562; EMBL:U25853; NID:g818908; PIDN:AAA85729.1; PI
 A:Experimental source: strain E71/M1
 A:Note: allele 2
 A:Accession: S77672
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-128 <KA2>
 A:Cross-references: UNIPARC:UPI0000000562; EMBL:U25854; NID:g818910; PIDN:AAA85730.1; PI
 A:Experimental source: strain E2/M3
 A:Note: allele 3
 A:Genetics:
 A:Gene: ska
 C:Superfamily: streptokinase
 C:Keywords: hydrolase; plasminogen activator; virulence
 Alignment Scores:
 Pred. No.: 3,14e-34 Length: 128
 Score: 620.00 Matches: 118
 Percent Similarity: 95.31% Conservative: 4
 Best Local Similarity: 92.19% Mismatches: 6
 Query Match: 16.64% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x S77671 (1-128)
 QY 1008 GTGGCGGTAGACCATATTAAGAAAAACCAATACAAACCAAGCGAATCTGTGATG 1067
 Db 1 ValAlaGValAlaGProTyrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 20
 QY 1068 GAATATATCTGACGATTACTTCCTTAAACCTGATGACGATTTGACACCGAGCTCAAA 1127
 Db 21 GluLysThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 40
 QY 1128 GATCTAGCTATTAAGAAACCTAGCTATCGGTGACACATCATCTCAAGAAATAC 1187
 Db 41 AspThrLysLeuLeuLysThrLeuAlaLeuGlyAspThrLysThrSerGlnGluLeu 60
 QY 1188 GCTCAAGACCAAGCATTTTAAACAAACCAACCGAGGCTATGATTAATGAAACGTGAC 1247
 Db 61 AlaGlnAlaGlnSerLysLeuAsnLysThrHisProGlyTyrThrLysGlnAlaGAsp 80

QY 1248 TCCTCAATCGTCACTGACATGACATGATTTCCGTAGATTTTACCAATGATGACAG 1307
 Db 81 SerSerLeuValThrHisAspAsnAspLysLeuHisGlnThrLysLeuProMetCysGlnGlu 100
 QY 1308 TTTACTTACCGGTGTTAAAAATCGGAAACAGCTTATGACATGATTAATAAAATCGTCTG 1367
 Db 101 PheThrTyGlnValLysAsnArgGlnGlnAlaTyGlnLysLeuProLysThrGlyLe 120
 QY 1368 AATGAGAAATTAACAACATGAC 1391
 Db 121 LysGluLysThrAsnAsnThrAsp 128
 RESULT 11
 S14428
 fibronectin precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
 C:Accession: S14428; S14455; A22319; S46203; S004559; A27252; I59049
 R:Hyne, R.O.
 Submitted to the EMBL Data Library, July 1989
 A:Reference number: S14428
 A:Accession: S14428
 A:Molecule type: mRNA
 A:Residues: 1-2477 <HYN>
 A:Cross-references: UNIPROT:P04937; UNIPARC:UPI000012ATC6; EMBL:X15906; NID:g56163; PID
 R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hyne, R.O.
 EMBL J. 6, 2573-2580, 1987
 A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
 A:Reference number: S12455; MUID:88054951; PMID:2445560
 A:Accession: S12455
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 609-1810,'T',1812-2283 <SCH>
 A:Cross-references: UNIPARC:UPI0000177AE; EMBL:X15906
 R:Tamkun, J.W.; Schwarzbauer, J.E.; Hyne, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
 A:Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing.
 A:Reference number: A22319; MUID:84298097; PMID:6089177
 A:Accession: A22319
 A:Molecule type: DNA
 A:Residues: 2052-2237 <TAM>
 A:Cross-references: UNIPARC:UPI0000177AEB
 R:Falkenberg, C.; Englund, J.J.; Thogersen, I.B.; Salvesen, G.; Akerstrom, B.
 Biochem. J. 301, 745-751, 1994
 A:Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in A:Reference number: S46203; MUID:94330948; PMID:7519849
 A:Accession: S46203
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1183-1192; GLN',1268,'P',1270-1271,'D',1273,'CF',1276,'PY',1385-1399 <FAL>
 A:Cross-references: UNIPARC:UPI0000177AEC; UNIPARC:UPI0000177AED; UNIPARC:UPI0000177AEE
 R:Patel, R.S.; Odehmat, E.; Schwarzbauer, J.E.; Hyne, R.O.
 EMBL J. 6, 2565-2572, 1987
 A:Title: Organization of the fibronectin gene provides evidence for exon shuffling during A:Reference number: S00459; MUID:88054950; PMID:3119323
 A:Accession: S00459
 A:Molecule type: DNA
 A:Residues: 1-139;2382-2477 <PAT>
 A:Cross-references: UNIPARC:UPI000017095C; UNIPARC:UPI0000177AEF; EMBL:X05831
 R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hyne, R.O.
 Cell 33, 421-431, 1983
 A:Title: Three different fibronectin mRNAs arise by alternative splicing within the cod A:Reference number: A27252; MUID:84082067; PMID:6317187
 A:Accession: A27252
 A:Molecule type: mRNA
 A:Residues: 1586-1720,'T',1722,1813-2477 <SC2>
 A:Cross-references: UNIPARC:UPI0000177AF0
 R:Odehmat, E.; Tamkun, J.W.; Hyne, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
 A:Title: Repeating modular structure of the fibronectin gene: Relationship to protein str A:Reference number: I59049; MUID:86016741; PMID:3863113
 A:Accession: I59049


```

QY 1186 TAGCTCAAGCACAAGACATTTAAACAAAACCAAGGCTTATGAGTTATGAACGTG 1245
    |||
    |||
Db 407 euValGlnThrArgly-----GlyAsn 415
QY 1246 ACTGCTCAATGCTCACTGATGCAATGACATTTTCGTACGATTTTACCAATGATCAAG 1305
    |||
    |||
Db 415 efaAnglyAlaIeuCysHis-----421
QY 1306 AGTTTACTTACCGGTGTAAAAATCCGGAAACAAGCTTATAGATCAATAAAAATCTGTC 1365
    |||
    |||
Db 422 --PheProPheIeuYrSerAenArg-----429
QY 1366 TGAATGAAGAAATAAACAACACTGACCTGATCTCTGAAGAAATTAAGTCTTAAAAAG 1425
    |||
    |||
Db 430 -----AenYrSerAenPcySerHisGly-----438
QY 1426 GGGAAAGCGGTATGATCCCTTTGATTCGAGTCACTTGA---CTGTTCACCAATCAAT 1482
    |||
    |||
Db 439 -----ArgArgAenAenMetYrTrpCysGlyThrThrGln 451
QY 1483 ACGTTGATGTCGATACCAAC-----GAATTCGTAAGAAAGAGGACGCTCTTA 1530
    |||
    |||
Db 451 enYrAenPAlaAenGlnYrPheCyeProMetAlaAlaHisGlnGlnIleCysT 471
QY 1531 CAGCTAGGCAACGTAACCTTAGACTTCAGAGATTATAGATCTCTGTAAGGCTTAAC 1590
    |||
    |||
Db 471 hrtThaAnglu-----474
QY 1591 TACCTTACACAAATCTCGATGCTTTTGGTATTATGACATTAACCTTAACGTAAGAAAGTNG 1650
    |||
    |||
Db 475 -----GlyValMet---TyrArgIleGlyAenGlnTrpA 485
QY 1651 AGGATATATCAGATGACACCAACCGTATCATACCGTTATATGCGCAAGCAGCCGAG 1710
    |||
    |||
Db 485 sPlvSGlnHisAenP-----LeuG 491
QY 1711 GAGAGATATGCTAGTACCAATTTAGTGTGTGTGCGCAGCGCA---CAGATTGTATCC- 1766
    |||
    |||
Db 491 LysHisMetMetAArgCysThrCysValGlyAenGlyArgGlyIleThrAlaCysGlnIleProt 511
QY 1767 -----ATAGCTGAGAAAGTGTTTTATGATCATGCTGCTGAGACTTCCTTAATGTGTGGAG 1818
    |||
    |||
Db 511 yrSerGlnLeuAArgAenGlnCysIle-----ValAenAenPAlaIleThrYrAenValAenA 529
QY 1819 AAAGTGGGAGAACCCCTACCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 1878
    |||
    |||
Db 529 sPlrPheHisIleYrGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 549
QY 1879 GCAGCGGAGCGATCACTTCACTTCTAGAAATAGATGACAAAGATGACAGACAGAGACAT 1938
    |||
    |||
Db 549 LysArgGlyArgTrpYrCysAenProIleAenPArgCysGlnAenPArgGlnIleArgThrP 569
QY 1939 CTTATAGATTTGAGACACTTGGACCAAGAGATATGACAGAAACCTGTCTCACTGCA 1998
    |||
    |||
Db 569 heYrGlnIleGlyAenPserTrp---GlnYrPheValHisGlyValArgYrGlnCysT 588
QY 1999 TCTGCAACGAGCAAGCGCGGAGAGAGTGTGAATGTGACAGAGCAACCTGTGACAGCA 2058
    |||
    |||
Db 588 yrCysYrGlyArgGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIle 605
QY 2059 CATGAGGAGGATCTGCGCC 2078
    |||
    |||
Db 605 yrProGlyThrThrGlyPro 611

```

```

A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77680
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <RAP>
A:Cross-references: UNIPROT:Q54687; UNIPARC:UPI00000BD2D3; EMBL:U25862; NID:g818926; PIR:
A:Experimental source: strain E751/M17
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 8.68e-11 Length: 128
Score: 569.00 Matches: 110
Percent Similarity: 91.41% Conservative: 7
Best Local Similarity: 85.94% Mismatch: 11
Query Match: 15.27% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x S77680 (1-128)

QY 1008 GTGCGGCTTACACCATTAAGAAACCAATACAAACCAAGCGAAATCTGTGATGTG 1067
    |||
    |||
Db 1 ValArgValArgProTrpYrGlnYrProIleGlnThrProAlaYrSerValAenPAl 20
QY 1068 GAATATATCTGTACAGTTACTTCCCTTAACCCCTGATGACGATTTTCAAGCCAGGCTCAAA 1127
    |||
    |||
Db 21 ArgGlyThrValGlnPheThrProLeuAenPArgAenPArgAenPArgAenPArgAenPArgAenPArg 40
QY 1128 GATATTAAGCTATTTGAAACCACTAGCTATCGGTGACACACACATGACATGCAAGATTACTA 1187
    |||
    |||
Db 41 AsnThrIleLeuLeuYrThrLeuAlaIleGlyGlyThrValIleThrSerGlnIleLeu 60
QY 1188 GCTCAAGCACAAGACATTTTAAACAAAACCAAGCGCTATAGATTTATGACGTGAC 1247
    |||
    |||
Db 61 AlAGlnAlaGlnSerIleLeuAenGlnSerHisPheAenPArgYrThrIleYrGlnArgA 80
QY 1248 TCTGCAATGCTCACTGATGCAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 1307
    |||
    |||
Db 81 SerSerIleValIleHisAenAenPAlaPheArgThrIleLeuProMetAenPArgGlnGln 100
QY 1308 TTTATTAACCGGTAAAAATCGGGAACAGCTTATGATGATCAATAAAAATCTGCTG 1367
    |||
    |||
Db 101 PheThrIleHisIleYrAenPArgGlnGlnAlaIleYrIleAenYrIleYrSerGlyGln 120
QY 1368 AATGAAGAAATTAACAACACTGAC 1391
    |||
    |||
Db 121 GlnGlnYrThrAenAenThrAenP 128

RESULT 13
S77688
streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77688
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whitlam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77688
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <RAP>
A:Cross-references: UNIPROT:Q54695; UNIPARC:UPI00000BD8DC; EMBL:U25870; NID:g818942; PIR:
A:Experimental source: strain E76/M72
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Alignment Scores:
Pred. No.: 5.61e-30 Length: 128

```

```

Score: 557.00 Matches: 107
Percent Similarity: 89.84% Conservative: 13
Best Local Similarity: 83.59% Mismatches: 8
Query Match: 14.95% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x S77688 (1-128)

QY 1008 GTGCGCGTTAGACATATTAAGAAAAACCAATACAAACGCGAAATCTGTGATGNG 1067
D 1 ValAGValAGProTyrLysGluYsProlIeGlnThProAlaYsSerValAptile 20
QY 1068 GAATATACGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGTCTCAA 1127
D 21 ArgTyrThrValGlnPheThrProLeuLenProAdaPAspPheYsProValLeuYs 40
QY 1128 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTA 1187
D 41 AspThrLysLeuLeuYsThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 60
QY 1188 GCTCAGACCAAGCATTTTAAACAAAACGACGCGTATACGATTTATGACGTCAC 1247
D 61 AlaGlnAlaGlnSerIleLeuIleGluSerHisProAspTyrThrIleTyrGluArgAsp 80
QY 1248 TCTCATCGTCACTCATGACATGACATTTTCGTCATGATTTTACCAATGATCAAGAG 1307
D 81 SerSerIleValThnHisAspAsnAspIlePheAlaGlnIleLeuProThrAspGln 100
QY 1308 TTTAACCTACCGTGTAAATAATCGGAGACAGCTTATAGATCAATAAATAATCTGCTG 1367
D 101 PheThrTyrHisValIysAsnArgGluGlnAlaTyrLysAlaAsnSerLysThrAspIle 120
QY 1368 AATGAGAAATTAACAACATCGAC 1391
D 121 LysGluYsThrAsnAsnThrAsp 128

RESULT 14
A43908
fibronectin - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
C/Accession: A43908
R/Description: D.W.; Norton, P.A.; Hynes, R.O.
Dev. Biol. 149, 357-369, 1992
A/Title: Identification and characterization of alternatively spliced fibronectin mRNAs
A/Reference number: A43908; MUID:92111942; PMID:1730390
A/Accession: A43908
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-2481 <DB>
A/Cross-references: UNIPARC:UPI0000177AE7; GB:M77820
A/Note: sequence extracted from NCBI Backbone (NCBI:P:77473)
C/Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C/Keywords: duplication; extracellular matrix; glycoprotein; heterodimer
F:55-90/Domain: fibronectin type I repeat homology <1F1>
F:100-138/Domain: fibronectin type I repeat homology <1F2>
F:144-182/Domain: fibronectin type I repeat homology <1F3>
F:189-228/Domain: fibronectin type I repeat homology <1F4>
F:234-273/Domain: fibronectin type I repeat homology <1F5>
F:309-343/Domain: fibronectin type I repeat homology <1F6>
F:361-402/Domain: fibronectin type II repeat homology <2F1>
F:421-462/Domain: fibronectin type II repeat homology <2F2>
F:471-509/Domain: fibronectin type I repeat homology <1F7>
F:519-556/Domain: fibronectin type I repeat homology <1F8>
F:562-600/Domain: fibronectin type I repeat homology <1F9>
F:610-693/Domain: fibronectin type III repeat homology <FN3A>
F:719-801/Domain: fibronectin type III repeat homology <FN3B>
F:810-881/Domain: fibronectin type III repeat homology <FN3C>
F:906-988/Domain: fibronectin type III repeat homology <FN3D>
F:996-1077/Domain: fibronectin type III repeat homology <FN3E>
F:1086-1165/Domain: fibronectin type III repeat homology <FN3F>
F:1173-1258/Domain: fibronectin type III repeat homology <FN3G>
F:1266-1349/Domain: fibronectin type III repeat homology <FN3H>

```

```

F:1357-1440/Domain: fibronectin type III repeat homology <FN3I>
F:1448-1530/Domain: fibronectin type III repeat homology <FN3J>
F:1538-1620/Domain: fibronectin type III repeat homology <FN3K>
F:1615-1617/Region: cell attachment (R-G-D) motif
F:1632-1714/Domain: fibronectin type III repeat homology <FN3L>
F:1722-1804/Domain: fibronectin type III repeat homology <FN3M>
F:1812-1894/Domain: fibronectin type III repeat homology <FN3N>
F:1904-1985/Domain: fibronectin type III repeat homology <FN3O>
F:1993-2075/Domain: fibronectin type III repeat homology <FN3P>
F:2197-2277/Domain: fibronectin type III repeat homology <FN3Q>
F:2301-2340/Domain: fibronectin type I repeat homology <1F10>
F:2346-2383/Domain: fibronectin type I repeat homology <1F11>
F:2390-2425/Domain: fibronectin type I repeat homology <1F12>
F:55-81, 79-90, 100-128, 126-138, 144-172, 170-182, 189-218, 216-228, 234-263, 261-273, 309-336, 333-343, 345-353, 355-363, 365-373, 375-383, 385-393, 395-403, 405-413, 415-423, 425-433, 435-443, 445-453, 455-463, 465-473, 475-483, 485-493, 495-503, 505-513, 515-523, 525-533, 535-543, 545-553, 555-563, 565-573, 575-583, 585-593, 595-603, 605-613, 615-623, 625-633, 635-643, 645-653, 655-663, 665-673, 675-683, 685-693, 695-703, 705-713, 715-723, 725-733, 735-743, 745-753, 755-763, 765-773, 775-783, 785-793, 795-803, 805-813, 815-823, 825-833, 835-843, 845-853, 855-863, 865-873, 875-883, 885-893, 895-903, 905-913, 915-923, 925-933, 935-943, 945-953, 955-963, 965-973, 975-983, 985-993, 995-1003, 1005-1013, 1015-1023, 1025-1033, 1035-1043, 1045-1053, 1055-1063, 1065-1073, 1075-1083, 1085-1093, 1095-1103, 1105-1113, 1115-1123, 1125-1133, 1135-1143, 1145-1153, 1155-1163, 1165-1173, 1175-1183, 1185-1193, 1195-1203, 1205-1213, 1215-1223, 1225-1233, 1235-1243, 1245-1253, 1255-1263, 1265-1273, 1275-1283, 1285-1293, 1295-1303, 1305-1313, 1315-1323, 1325-1333, 1335-1343, 1345-1353, 1355-1363, 1365-1373, 1375-1383, 1385-1393, 1395-1403, 1405-1413, 1415-1423, 1425-1433, 1435-1443, 1445-1453, 1455-1463, 1465-1473, 1475-1483, 1485-1493, 1495-1503, 1505-1513, 1515-1523, 1525-1533, 1535-1543, 1545-1553, 1555-1563, 1565-1573, 1575-1583, 1585-1593, 1595-1603, 1605-1613, 1615-1623, 1625-1633, 1635-1643, 1645-1653, 1655-1663, 1665-1673, 1675-1683, 1685-1693, 1695-1703, 1705-1713, 1715-1723, 1725-1733, 1735-1743, 1745-1753, 1755-1763, 1765-1773, 1775-1783, 1785-1793, 1795-1803, 1805-1813, 1815-1823, 1825-1833, 1835-1843, 1845-1853, 1855-1863, 1865-1873, 1875-1883, 1885-1893, 1895-1903, 1905-1913, 1915-1923, 1925-1933, 1935-1943, 1945-1953, 1955-1963, 1965-1973, 1975-1983, 1985-1993, 1995-2003, 2005-2013, 2015-2023, 2025-2033, 2035-2043, 2045-2053, 2055-2063, 2065-2073, 2075-2083, 2085-2093, 2095-2103, 2105-2113, 2115-2123, 2125-2133, 2135-2143, 2145-2153, 2155-2163, 2165-2173, 2175-2183, 2185-2193, 2195-2203, 2205-2213, 2215-2223, 2225-2233, 2235-2243, 2245-2253, 2255-2263, 2265-2273, 2275-2283, 2285-2293, 2295-2303, 2305-2313, 2315-2323, 2325-2333, 2335-2343, 2345-2353, 2355-2363, 2365-2373, 2375-2383, 2385-2393, 2395-2403, 2405-2413, 2415-2423, 2425-2433, 2435-2443, 2445-2453, 2455-2463, 2465-2473, 2475-2483, 2485-2493, 2495-2503, 2505-2513, 2515-2523, 2525-2533, 2535-2543, 2545-2553, 2555-2563, 2565-2573, 2575-2583, 2585-2593, 2595-2603, 2605-2613, 2615-2623, 2625-2633, 2635-2643, 2645-2653, 2655-2663, 2665-2673, 2675-2683, 2685-2693, 2695-2703, 2705-2713, 2715-2723, 2725-2733, 2735-2743, 2745-2753, 2755-2763, 2765-2773, 2775-2783, 2785-2793, 2795-2803, 2805-2813, 2815-2823, 2825-2833, 2835-2843, 2845-2853, 2855-2863, 2865-2873, 2875-2883, 2885-2893, 2895-2903, 2905-2913, 2915-2923, 2925-2933, 2935-2943, 2945-2953, 2955-2963, 2965-2973, 2975-2983, 2985-2993, 2995-3003, 3005-3013, 3015-3023, 3025-3033, 3035-3043, 3045-3053, 3055-3063, 3065-3073, 3075-3083, 3085-3093, 3095-3103, 3105-3113, 3115-3123, 3125-3133, 3135-3143, 3145-3153, 3155-3163, 3165-3173, 3175-3183, 3185-3193, 3195-3203, 3205-3213, 3215-3223, 3225-3233, 3235-3243, 3245-3253, 3255-3263, 3265-3273, 3275-3283, 3285-3293, 3295-3303, 3305-3313, 3315-3323, 3325-3333, 3335-3343, 3345-3353, 3355-3363, 3365-3373, 3375-3383, 3385-3393, 3395-3403, 3405-3413, 3415-3423, 3425-3433, 3435-3443, 3445-3453, 3455-3463, 3465-3473, 3475-3483, 3485-3493, 3495-3503, 3505-3513, 3515-3523, 3525-3533, 3535-3543, 3545-3553, 3555-3563, 3565-3573, 3575-3583, 3585-3593, 3595-3603, 3605-3613, 3615-3623, 3625-3633, 3635-3643, 3645-3653, 3655-3663, 3665-3673, 3675-3683, 3685-3693, 3695-3703, 3705-3713, 3715-3723, 3725-3733, 3735-3743, 3745-3753, 3755-3763, 3765-3773, 3775-3783, 3785-3793, 3795-3803, 3805-3813, 3815-3823, 3825-3833, 3835-3843, 3845-3853, 3855-3863, 3865-3873, 3875-3883, 3885-3893, 3895-3903, 3905-3913, 3915-3923, 3925-3933, 3935-3943, 3945-3953, 3955-3963, 3965-3973, 3975-3983, 3985-3993, 3995-4003, 4005-4013, 4015-4023, 4025-4033, 4035-4043, 4045-4053, 4055-4063, 4065-4073, 4075-4083, 4085-4093, 4095-4103, 4105-4113, 4115-4123, 4125-4133, 4135-4143, 4145-4153, 4155-4163, 4165-4173, 4175-4183, 4185-4193, 4195-4203, 4205-4213, 4215-4223, 4225-4233, 4235-4243, 4245-4253, 4255-4263, 4265-4273, 4275-4283, 4285-4293, 4295-4303, 4305-4313, 4315-4323, 4325-4333, 4335-4343, 4345-4353, 4355-4363, 4365-4373, 4375-4383, 4385-4393, 4395-4403, 4405-4413, 4415-4423, 4425-4433, 4435-4443, 4445-4453, 4455-4463, 4465-4473, 4475-4483, 4485-4493, 4495-4503, 4505-4513, 4515-4523, 4525-4533, 4535-4543, 4545-4553, 4555-4563, 4565-4573, 4575-4583, 4585-4593, 4595-4603, 4605-4613, 4615-4623, 4625-4633, 4635-4643, 4645-4653, 4655-4663, 4665-4673, 4675-4683, 4685-4693, 4695-4703, 4705-4713, 4715-4723, 4725-4733, 4735-4743, 4745-4753, 4755-4763, 4765-4773, 4775-4783, 4785-4793, 4795-4803, 4805-4813, 4815-4823, 4825-4833, 4835-4843, 4845-4853, 4855-4863, 4865-4873, 4875-4883, 4885-4893, 4895-4903, 4905-4913, 4915-4923, 4925-4933, 4935-4943, 4945-4953, 4955-4963, 4965-4973, 4975-4983, 4985-4993, 4995-5003, 5005-5013, 5015-5023, 5025-5033, 5035-5043, 5045-5053, 5055-5063, 5065-5073, 5075-5083, 5085-5093, 5095-5103, 5105-5113, 5115-5123, 5125-5133, 5135-5143, 5145-5153, 5155-5163, 5165-5173, 5175-5183, 5185-5193, 5195-5203, 5205-5213, 5215-5223, 5225-5233, 5235-5243, 5245-5253, 5255-5263, 5265-5273, 5275-5283, 5285-5293, 5295-5303, 5305-5313, 5315-5323, 5325-5333, 5335-5343, 5345-5353, 5355-5363, 5365-5373, 5375-5383, 5385-5393, 5395-5403, 5405-5413, 5415-5423, 5425-5433, 5435-5443, 5445-5453, 5455-5463, 5465-5473, 5475-5483, 5485-5493, 5495-5503, 5505-5513, 5515-5523, 5525-5533, 5535-5543, 5545-5553, 5555-5563, 5565-5573, 5575-5583, 5585-5593, 5595-5603, 5605-5613, 5615-5623, 5625-5633, 5635-5643, 5645-5653, 5655-5663, 5665-5673, 5675-5683, 5685-5693, 5695-5703, 5705-5713, 5715-5723, 5725-5733, 5735-5743, 5745-5753, 5755-5763, 5765-5773, 5775-5783, 5785-5793, 5795-5803, 5805-5813, 5815-5823, 5825-5833, 5835-5843, 5845-5853, 5855-5863, 5865-5873, 5875-5883, 5885-5893, 5895-5903, 5905-5913, 5915-5923, 5925-5933, 5935-5943, 5945-5953, 5955-5963, 5965-5973, 5975-5983, 5985-5993, 5995-6003, 6005-6013, 6015-6023, 6025-6033, 6035-6043, 6045-6053, 6055-6063, 6065-6073, 6075-6083, 6085-6093, 6095-6103, 6105-6113, 6115-6123, 6125-6133, 6135-6143, 6145-6153, 6155-6163, 6165-6173, 6175-6183, 6185-6193, 6195-6203, 6205-6213, 6215-6223, 6225-6233, 6235-6243, 6245-6253, 6255-6263, 6265-6273, 6275-6283, 6285-6293, 6295-6303, 6305-6313, 6315-6323, 6325-6333, 6335-6343, 6345-6353, 6355-6363, 6365-6373, 6375-6383, 6385-6393, 6395-6403, 6405-6413, 6415-6423, 6425-6433, 6435-6443, 6445-6453, 6455-6463, 6465-6473, 6475-6483, 6485-6493, 6495-6503, 6505-6513, 6515-6523, 6525-6533, 6535-6543, 6545-6553, 6555-6563, 6565-6573, 6575-6583, 6585-6593, 6595-6603, 6605-6613, 6615-6623, 6625-6633, 6635-6643, 6645-6653, 6655-6663, 6665-6673, 6675-6683, 6685-6693, 6695-6703, 6705-6713, 6715-6723, 6725-6733, 6735-6743, 6745-6753, 6755-6763, 6765-6773, 6775-6783, 6785-6793, 6795-6803, 6805-6813, 6815-6823, 6825-6833, 6835-6843, 6845-6853, 6855-6863, 6865-6873, 6875-6883, 6885-6893, 6895-6903, 6905-6913, 6915-6923, 6925-6933, 6935-6943, 6945-6953, 6955-6963, 6965-6973, 6975-6983, 6985-6993, 6995-7003, 7005-7013, 7015-7023, 7025-7033, 7035-7043, 7045-7053, 7055-7063, 7065-7073, 7075-7083, 7085-7093, 7095-7103, 7105-7113, 7115-7123, 7125-7133, 7135-7143, 7145-7153, 7155-7163, 7165-7173, 7175-7183, 7185-7193, 7195-7203, 7205-7213, 7215-7223, 7225-7233, 7235-7243, 7245-7253, 7255-7263, 7265-7273, 7275-7283, 7285-7293, 7295-7303, 7305-7313, 7315-7323, 7325-7333, 7335-7343, 7345-7353, 7355-7363, 7365-7373, 7375-7383, 7385-7393, 7395-7403, 7405-7413, 7415-7423, 7425-7433, 7435-7443, 7445-7453, 7455-7463, 7465-7473, 7475-7483, 7485-7493, 7495-7503, 7505-7513, 7515-7523, 7525-7533, 7535-7543, 7545-7553, 7555-7563, 7565-7573, 7575-7583, 7585-7593, 7595-7603, 7605-7613, 7615-7623, 7625-7633, 7635-7643, 7645-7653, 7655-7663, 7665-7673, 7675-7683, 7685-7693, 7695-7703, 7705-7713, 7715-7723, 7725-7733, 7735-7743, 7745-7753, 7755-7763, 7765-7773, 7775-7783, 7785-7793, 7795-7803, 7805-7813, 7815-7823, 7825-7833, 7835-7843, 7845-7853, 7855-7863, 7865-7873, 7875-7883, 7885-7893, 7895-7903, 7905-7913, 7915-7923, 7925-7933, 7935-7943, 7945-7953, 7955-7963, 7965-7973, 7975-7983, 7985-7993, 7995-8003, 8005-8013, 8015-8023, 8025-8033, 8035-8043, 8045-8053, 8055-8063, 8065-8073, 8075-8083, 8085-8093, 8095-8103, 8105-8113, 8115-8123, 8125-8133, 8135-8143, 8145-8153, 8155-8163, 8165-8173, 8175-8183, 8185-8193, 8195-8203, 8205-8213, 8215-8223, 8225-8233, 8235-8243, 8245-8253, 8255-8263, 8265-8273, 8275-8283, 8285-8293, 8295-8303, 8305-8313, 8315-8323, 8325-8333, 8335-8343, 8345-8353, 8355-8363, 8365-8373, 8375-8383, 8385-8393, 8395-8403, 8405-8413, 8415-8423, 8425-8433, 8435-8443, 8445-8453, 8455-8463, 8465-8473, 8475-8483, 8485-8493, 8495-8503, 8505-8513, 8515-8523, 8525-8533, 8535-8543, 8545-8553, 8555-8563, 8565-8573, 8575-8583, 8585-8593, 8595-8603, 8605-8613, 8615-8623, 8625-8633, 8635-8643, 8645-8653, 8655-8663, 8665-8673, 8675-8683, 8685-8693, 8695-8703, 8705-8713, 8715-8723, 8725-8733, 8735-8743, 8745-8753, 8755-8763, 8765-8773, 8775-8783, 8785-8793, 8795-8803, 8805-8813, 8815-8823, 8825-8833, 8835-8843, 8845-8853, 8855-8863, 8865-8873, 8875-8883, 8885-8893, 8895-8903, 8905-8913, 8915-8923, 8925-8933, 8935-8943, 8945-8953, 8955-8963, 8965-8973, 8975-8983, 8985-8993, 8995-9003, 9005-9013, 9015-9023, 9025-9033, 9035-9043, 9045-9053, 9055-9063, 9065-9073, 9075-9083, 9085-9093, 9095-9103, 9105-9113, 9115-9123, 9125-9133, 9135-9143, 9145-9153, 9155-9163, 9165-9173, 9175-9183, 9185-9193, 9195-9203, 9205-9213, 9215-9223, 9225-9233, 9235-9243, 9245-9253, 9255-9263, 9265-9273, 9275-9283, 9285-9293, 9295-9303, 9305-9313, 9315-9323, 9325-9333, 9335-9343, 9345-9353, 9355-9363, 9365-9373, 9375-9383, 9385-9393, 9395-9403, 9405-9413, 9415-9423, 9425-9433, 9435-9443, 9445-9453, 9455-9463, 9465-9473, 9475-9483, 9485-9493, 9495-9503, 9505-9513, 9515-9523, 9525-9533, 9535-9543, 9545-9553, 9555-9563, 9565-9573, 9575-9583, 9585-9593, 9595-9603, 9605-9613, 9615-9623, 9625-9633, 9635-9643
```

```

Db      296 -----|||:::|||||:::|||||:::|
          -LeuTyrGlnPro-----AspSerGlnLeu-----
Qy      658 GTACTGTGAGGGAGCGAATCAAGACATTAGTCTTAATTTTGAATGATCTAACAT 717
Db      303 -----
Qy      718 CAGACCTGCTCATGAGAGAAAGACAGAGCGTTAAGTCCAAATCAAAACATTGG 777
Db      304 -----GlnProTyrG 307
Qy      778 CTACTGATAGTGCGCGAGTGTACATTAACCTTGAGAAAGCTGACTTACTAAGGCTATTC 837
Db      307 1y-----
Qy      838 AAGAACAAATTGATCGCTAACGTCACAGTAAGACGACTTGGAGTGATGATTGG 897
Db      307 -----
Qy      898 CAAGCGATGCAACATTACTGATCGAAGCGCAAGGTCTACTTGTGTCAGAACAGATGTT 957
Db      308 -----HisCysValThrAsp---AsnGlyValLeuTyrSerLeu-----GlyM 321
Qy      958 CGGTAACTTGCCCGACCACTGTCCAGAAATTTTGTAAAGCGGACATGTCGGCGTTA 1017
Db      321 eCArgTTrpLeuArgThrGlnGlySerTyrGlnMetLeuCysThr----- 335
Qy      1018 GACCATATTAAGAAAAACCAATACAAACCAAGCGAAATCTGTGATGTGAATATATCTG 1077
Db      335 -----
Qy      1078 TACAGTTACTCCCTTAAACCTGATGACGATTTCAAGCAGGTCCTCAAGATACTAAGC 1137
Db      335 -----
Qy      1138 TATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGATTACTAGCTCAAGCAC 1197
Db      336 -----CysLeuGlyAsnGlyValSerCysGlnGlnThrValAlaThrIleT 351
Qy      1198 AAGACATTTTAAACAAAAC-----CACCCAGGCTATA 1230
Db      351 hrPheGlyGlyAsnAlaAsnGlyLysProCysAlaIleProPheThrHisAspGlyLysT 371
Qy      1231 CGATTATGACGTGACTCTCTCA-----ATGTCACATC 1263
Db      371 hrTyrTyrSerCysThrSerGlnGlyArgGlnAspGlyLysLeuTyrCysAlaThrThrS 391
Qy      1264 ATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGATTACTTACTCGTGT 1323
Db      391 eAsnTyrAsp-----SerAspLysLysTyrSerPheCysThr- 403
Qy      1334 AAAATCGGACACGCTTATAGATCAATAAATCTGGT----- 1364
Db      404 -----GlnGlnLeuAlaLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuc 421
Qy      1365 -----CTGATGAAGAATAAACAACCTGACCTGATCTCGAATAATATT 1410
Db      421 yAsnProPheProPheLeuTyrAsnAsnLeuAsnTyrThrAspCysThrSerGln----- 438
Qy      1411 ACGTCCTTAAAAAGGGGAAAAAGCCGTATGATCCCTTGTATGCCAGATCTTGAATGTT 1470
Db      439 -----GlyArgGlnAsp----- 442
Qy      1471 TCACCATCAATACGTTGATGTGATGATCAAC----- 1502
Db      443 --SerMetLysThrCysGlyThrThrAlaSerTyrAspAlaAspGlnLysPheGlyPheC 462
Qy      1503 --GAATGTCTAAAGTGAAGCGAGCTCTTAAACAGCTGACGAGTAAGTAACTTACAGAG 1560
Db      462 ySPromerAlaAlaHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 475
Qy      1561 ATTATAGATCTCTGTGATTAAGGCTAAACTACTACTACAAACATCTCGATCTTTGGTA 1620
          |||:

```

```

Db      476 -----GlyV 477
Qy      1621 TTATGACTATACCTTAATCTGAAAAAGTAGAGATTAATCAGATGACACCAACCGTATCA 1680
          :|||:::|:::|:::|:::|:::|
Db      477 alMet---TyrArgValGlyAspGlnTrpAspLysGlnHisAspGlnGlnHisMetMetC 496
Qy      1681 TAAACGTTTATATGAGGCAAG---CGACCCGAGAGAGAAAGTACTAGTCAATTTAGCTG 1737
Db      496 rGysThrCysValGlyAsnGlyArgGlyGlnTyrSerCysValAlaTyrSer----- 513
Qy      1738 GTGCTGCGCAGCGCCACACAGATTGTACCATAGCTGAGAAAGTGTTTATCATGCTGCTG 1797
Db      514 -----GlnLeuLysAspGlnCysIle-----ValAspG 523
Qy      1798 GAACCTCTATAGTGTCGAGAAACGTCGAGAAAGCCCTACCAAGCGCTGATGATGCTAG 1857
Db      523 1yLeuThrTyrAsnValAsnSerSerPheThrLysLeuHisGlnGlnGlnHisMetMetC 543
Qy      1858 ATTGTACTTGTGCTGGGAGAAAGCGAGCCGACCATCACTTGCATTTGTAATAGATGCA 1917
Db      543 snCysThrCysPheGlyGlnGlyArgGlyArgTyrCysAspAlaIleAspGlnCysG 563
Qy      1918 ACGATCAGACACAAAGACATCTATAGAAATTGAGACACTGAGACAAAGAGATATATC 1977
Db      563 1AspThrGlnThrArgGlnPheTyrGlnIleGlyAspSerTrp---GlnLysHisLeuG 582
Qy      1978 GAGGAACCTGCTCCAGTCATCTGCACAGCAACGCGGAGAGTGAAGTGTGGA 2037
Db      582 1nglyValGlnTyrGlnCysTyrCysTyrGlyLysGlyIleGlyGlnTrpHisCysGln- 601
Qy      2038 GGCACACCTGCTGCACACATCGAGCGGATGCGCCC 2078
Db      602 -----ProLeuSerThrSerGlnAlaGlyThrGlyPro 612

```

RESULT 15

```

S77679
Streptokinase A (BC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77679
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.J.; Whitlam, T.S.; Sawyer, S.A.; Musser
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:56037795; PMID:7565111
A:Accession: S77679
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54686; UNIPARC:UPI000008D6F; EMBL:U25861; NID:9818924; PIR
A:Experimental source: strain ET50/M43
C:GeneticS:
A:Gene: ska
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

```

Alignment Scores:

```

Pred. No.: 1,22e-29 Length: 128
Score: 552.00 Matches: 106
Percent Similarity: 89.06% Conservative: 8
Best local Similarity: 82.81% Mismatches: 14
Query Match: 14.81% Indels: 0
DB: 2 Gaps: 0

```

US-09-940-235-12 (1-2096) x S77679 (1-128)

```

Qy      1008 GTGCGCGCTGACCATTAAGAAACCAATACAAACCAAGGAATCTGTGATGTG 1067
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1 ValArgValArgProTyrLysGlnLysProIleGlnThrProAlaLysSerValAspIle 20
Qy      1068 GAATATATCTGTATACCTTACTCCCTTAAACCGTATGACGATTTCAAGACAGTCTCAAA 1127
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      21 ArgTyrValValGlnPheThrProLeuAsnProAspAspPheThrProValLeuLys 40

```

```

Qy 1128 GATACCTAAGCTATTGAAAACTAGCTATCGGTGACACCATCAATCTCAAGATTACTA 1187
    |||||
Db 41 AsptThrIysLeuLeuIysThrIleuAlaIleGIYAspThrIleThrSerGIuLeuLeu 60
    |||||
Qy 1188 GCTCAGACCAAAAGCATTTTAAACAAAACCAACCGCTATACGATTATGAAAGGTGAC 1247
    |||||
Db 61 AlaGlnAlaGlnSerIleLeuIleGIuSerHisProAsnIyrThrIleHisGIuArGAsp 80
    |||||
Qy 1248 TCCGCAATCGTCACTCATGACAAATGACATTTCCGTACGATTTTACCAATGATCAAGAG 1307
    |||||
Db 81 SerSerIleValThrHisAspAsnAspIlePheArGThrIleLeuProThrAspGIuGIu 100
    |||||
Qy 1308 TTACTTACCGGTGTTAAATAATCGGGAACAAGCTTATAGGATCAATATAATAATCTGCTG 1367
    |||||
Db 101 PheThrIyrHisValIlyAsnArGIuGlnAlaIyrIlyAlaAsnSerIyrStnGIyIle 120
    |||||
Qy 1368 AATGAGAAATTAACACACTGAC 1391
    |||||
Db 121 LysGIuLysThrAsnAsnThrAsp 128
    |||||

```

Search completed: January 28, 2006, 02:38:18
 Job time : 56.1209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 28, 2006, 01:48:09 ; Search time 225.882 Seconds
(without alignments)
13093.422 Million cell updates/sec

Title: US-09-940-235-12

Perfect score: 3726

Sequence: 1 cgaagaccatcatgttgcgtt.....ccttcacgatgtcgttag 2096

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
MODEL=frame_n2p.model -DEV=xlp
-O=/cgn2.1/USPTO_spool_p/US09940235/runat_27012006_144218_27563/app_query.fasta_1.7708
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235 @CGN_1.1.1355 @runat_27012006_144218_27563 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

1: UniProt_05.80.*
2: uniprot_sprot.*
3: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1984	53.2	440	1	STRP_STREQ
2	1985	52.5	414	2	Q53284_STREQ
3	1943	52.1	436	2	Q6UK57_STREQ
4	1942	52.1	440	1	STRP_STRS1
5	1814	48.7	440	2	Q7X0Y1_STRPY
6	1802	48.4	440	2	Q7X0Y8_STRPY
7	1802	48.4	440	2	Q8K5R8_STRP3
8	1794	48.1	440	1	STRP_STRPY
9	1794	48.1	440	2	Q532X6_STRPY
10	1787	48.0	440	2	Q7X0Y2_STRPY
11	1758	47.2	440	2	Q7X0Y7_STRPY
12	1758	47.2	440	2	Q5X9T6_STRP6
13	1755	47.1	440	2	Q8NZ46_STRP8
14	1723	46.2	440	2	Q7X0Y3_STRPY
15	1715	46.0	440	2	Q7X0Y0_STRPY
16	1715	46.0	440	2	Q7X0Y5_STRPY

17	1708	45.8	440	2	Q7X0X8_STRPY	Q7X0X8 streptococc
18	1686	45.2	440	2	Q7X0X9_STRPY	Q7X0X9 streptococc
19	1677	45.0	440	2	Q7X0X7_STRPY	Q7X0X7 streptococc
20	1670	44.8	440	2	Q57391_STRPY	Q57391 streptococc
21	1657	44.5	440	2	Q7X0Y6_STRPY	Q7X0Y6 streptococc
22	1652	44.3	440	1	STRP_STRPY	Q7X0Y6 streptococc
23	1624	43.6	432	2	Q7X0Y4_STRPY	Q7X0Y4 streptococc
24	728	19.5	141	2	Q7X0X2_STREQ	Q7X0X2 streptococc
25	727	19.5	141	2	Q7X0X3_STREQ	Q7X0X3 streptococc
26	724	19.4	141	2	Q7X0M1_STREQ	Q7X0M1 streptococc
27	723	19.4	141	2	Q7X0S7_STREQ	Q7X0S7 streptococc
28	717	19.2	141	2	Q7X0M4_STREQ	Q7X0M4 streptococc
29	717	19.2	141	2	Q7X0M3_STREQ	Q7X0M3 streptococc
30	715	19.2	141	2	Q7X0M5_STREQ	Q7X0M5 streptococc
31	714	19.2	141	2	Q7X0X5_STREQ	Q7X0X5 streptococc
32	712	19.1	141	2	Q7X0X6_STREQ	Q7X0X6 streptococc
33	711	19.1	141	2	Q7X0M2_STREQ	Q7X0M2 streptococc
34	711	19.1	141	2	Q7X0M9_STREQ	Q7X0M9 streptococc
35	710	19.1	141	2	Q7X0M6_STREQ	Q7X0M6 streptococc
36	707	19.0	141	2	Q7X0M0_STREQ	Q7X0M0 streptococc
37	705	18.9	141	2	Q7X0V5_STRPY	Q7X0V5 streptococc
38	704	18.9	141	2	Q7X0X0_STREQ	Q7X0X0 streptococc
39	701	18.8	141	2	Q7X0W7_STREQ	Q7X0W7 streptococc
40	697	18.7	141	2	Q7X0M4_STREQ	Q7X0M4 streptococc
41	696	18.7	141	2	Q7X0M8_STREQ	Q7X0M8 streptococc
42	693	18.6	141	2	Q7X0X1_STREQ	Q7X0X1 streptococc
43	690	18.5	141	2	Q7X0R2_STRPY	Q7X0R2 streptococc
44	686	18.4	141	2	Q7X0S4_STRPY	Q7X0S4 streptococc
45	684	18.4	141	2	Q7X0S0_STRPY	Q7X0S0 streptococc

ALIGNMENTS

```
RESULT 1
STRP_STREQ STANDARD; PRT; 440 AA.
ID P00779;
AC 21-JUN-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Streptokinase C precursor.
GN Name=ekc;
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H46A;
RX MEDLINE=85232082; PubMed=2989113; DOI=10.1016/0378-1119(85)90145-3;
RA Walke H., Roe B.A., Ferretti J.U.;
RT "Nucleotide sequence of the streptokinase gene from Streptococcus
equisimilis H46A."
RL Gene 34:357-362(1985).
RN [2]
RP PROTEIN SEQUENCE OF 27-440.
RX MEDLINE=83127125; PubMed=6760891;
RA Jackson K.W., Tang J.;
RT "Complete amino acid sequence of streptokinase and its homology with
serine proteases."
RL Biochemistry 21:6620-6625(1982).
-!- FUNCTION: This protein is not a protease, but it activates
plasminogen by complexing with it. As a potential virulence
factor, it is thought to prevent the formation of effective fibrin
barriers around the site of infection, thereby contributing to the
invasiveness of the cells.
-!- Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
```



```
DR EMBL; K02986; AAA26974.1; -; Genomic DNA.
DR EMBL; X72832; CAA51351.1; -; Genomic DNA.
DR PIR; A00967; BZSO.
DR PIR; A22801; A22801.
DR PDB; 1BML; X-ray; C/D=38-399.
DR PDB; 1L4D; X-ray; B=40-173.
DR PDB; 1L4Z; X-ray; B=27-173.
DR PDB; 1QOR; X-ray; A/B/C/D=177-314.
DR SMR; P00779; 38-398.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
DR 3D-structure; Direct protein sequencing; Plasminogen activation;
KW Signal; Variance.
FT SIGNAL 1 26
FT CHAIN 1 27
FT VARIANT 195 195
FT VARIANT 207 207
FT CONFLICT 298 300
FT CONFLICT 438 438
FT STRAND 180 180
FT STRAND 184 194
FT TURN 199 200
FT TURN 203 204
FT STRAND 205 205
FT STRAND 209 214
FT TURN 216 217
FT STRAND 219 221
FT HELIX 222 236
FT TURN 238 239
FT STRAND 240 252
FT TURN 253 254
FT STRAND 259 260
FT STRAND 268 270
FT STRAND 275 275
FT STRAND 278 280
FT TURN 282 284
FT STRAND 287 289
FT STRAND 292 304
FT TURN 305 306
SQ SEQUENCE 440 AA; 50140 MW; 8FC1F22648AC77A CRC64;

Alignment Scores:
Pred. No.: 2.86e-132 Length: 440
Score: 1984.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 53.25% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-12 (1-2096) x STRP_STREQ (1-440)
QY 588 ATTGCTGACCTGAGTGGCTGTAGACGCTCATCTGTCAACAAGCCAAATTGTTT 647
DB 27 TLeaIaGlyProGILtTPrLeuLeuAspArgProSeValAsnAsnSerGlnLeuVal 46
QY 648 AGCGTGTCTGTACTGTGAGGGAAGATCAAGACATTAGTCTTAATTTTGAATC 707
DB 47 SeValaIaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuysPhePheGluIle 66
QY 708 GATCTAATCATCAGCAGCTGCTCATGAGGAGAAAGACAGACAGCGTTAAGTCCAAATCA 767
DB 67 AspLeuThrSerArgProAlaHisGlyGlyLysThrGlnGlnGlyLeuSerProLysSer 86
QY 768 AAACCATTTGCTACTGATAGTGGCGGAGTGCACATAAATTGAGAAAGCTGACTTA 827
DB 87 LysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlnLysAlaAspLeu 106
QY 828 AAGCTATTCAAGAACAAATTGATCGCTAAGTCCACAGTAAAGACGACTTACTTGAAGTC 887
DB 107 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluVal 126
```

```
QY 888 ATTGATTTTGCAGGACGATGCAACCATTAAGTATGATGAAACGCAAGGTCTAATTGTGTAC 947
DB 127 TLeaSPheAlaSerAspAlaThrIleThrAspArgAsnLysValTyrPheAlaAsp 146
QY 948 AAAGATGTTTGGTAACTTTGCCGACCAACCTGTCCAGAAATTTTGTAAAGCGACAT 1007
DB 147 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 166
QY 1008 GTCCGCGTTGACCATTTAAAGAAAAACCAATCAAAACCAAGCGAAATCTGTGATGTG 1067
DB 167 ValArgValArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 186
QY 1068 GAATATATGCTGATGATTATCTCCCTTAAACCGATGATGACATTTCAACAGGCTCAAA 1127
DB 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1128 GATATCAAGCTAATTTGAAAAACATAGCTATGCGTACACATCAATCTCAAGAAATTA 1187
DB 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeu 226
QY 1188 GCTCAAGCACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGAC 1247
DB 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 246
QY 1248 TCCTCAATGCTGACCTATGACATTTGCAATTTCCGATGATTTTAAACCAATGATCAAG 1307
DB 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
QY 1308 TTTACTTACCGTGTAAATTCGGGAACAGCTTATAGATCAATAAAAAATCGTGTCTG 1367
DB 267 PheThrTyrArgValLysAsnArgGlnGlnAlaIleTyrArgLysAsnLysSerGlyLeu 286
QY 1368 AATGAAGAAATTAACAACTGACCTGATGCTGAGCAATTTACGCTTTAAAAAGGG 1427
DB 287 AsnGlnGlnLysAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
QY 1428 GAAAGCCGTATGATCCCTTGTATGCGACAGTCACTTGAAACGTTCACATCAATAATAG 1487
DB 307 GlnLysProTyrAspProPheAspArgSerHisLysLeuLysLeuPheThrIleLysTyrVal 326
QY 1488 GATGTGATPACCAACGAATTTGTAAAAAGTGAACAGCTCTTAAACAGCTAAGCAAGTAC 1547
DB 327 AspValAspThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysAsn 346
QY 1548 TTAGACTTCAAGATTTATTCAGTCTCTGTATAGGCTTAACCTTCAACAATCTTC 1607
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1608 GATGCTTTTGGTATTTATGACCTTACTTAACTGGAAGAGTAAAGATATATCAGATGAC 1667
DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspAsp 386
QY 1668 ACCAACCGTATCAATACCGTTTATATGGCGCAAGCCGAGAGAGAGAAATGCTAGTAC 1727
DB 387 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlnLysAlaSerTyr 406
QY 1728 CATTAGCT 1736
DB 407 HisLeuAla 409

RESULT 2
Q53284 STREQ
ID Q53284 STREQ PRELIMINARY; PRT; 414 AA.
AC Q53284;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE SKC-2.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
```

RP NUCLEOTIDE SEQUENCE.
 RA Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,
 RA Rubiera R., Pedraza A., Padron G., Antuch W., la Fuente J., et al.,
 RL "High level expression of streptokinase in *Escherichia coli*.";
 DR Biotechnology 0:1138-1142 (1992).
 DR EMBL: S46536; AAC60418.1; -; Genomic_DNA.
 DR PDB: 1C4P; X-ray; A/B/C/D=149-285.
 DR SMR: Q53284; 12-372.
 DR GO: GO:0008243; F:Plasminogen activator activity; IEA.
 DR GO: GO:0005515; F:Protein binding; IEA.
 DR InterPro: IPR004093; Streptokinase.
 DR InterPro: IPR008124; Streptokinase.
 DR Pfam: PF02821; Streptokinase; 3.
 DR PRINTS: PR01753; STREPKINASE.
 DR SEQUENCE 414 AA; 47254 MW; F75BE5831B766904 CRC64;

Alignment Scores:
 Pred. No.: 3,256-130 Length: 414
 Score: 1955.00 Matches: 377
 Percent Similarity: 98.96% Conservative: 2
 Best Local Similarity: 98.43% Mismatches: 4
 Query Match: 52.47% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q53284_STREQ (1-414)

QY 588 ATTGCTGAGCTGAGTGGCTGAGACCGTCCATCTGTCAACAACGCAATGTTGTT 647
 Db 1 MetAlaGlyProGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 20
 QY 648 AGCGTTGCTGCTACTGTGTTGAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
 Db 21 SerValAlaGlyThrValGlnGlyThrAsnGlnAspIleSerLeuysPheGluIle 40
 QY 708 GATTAAACATCAACGACCTCTCTCAATGAGAGAAAGACAGACGCTTAAGTCCAAATCA 767
 Db 41 AspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGlyLeuSerProLysSer 60
 QY 768 AAACGATTGCTACTGATGTGGCGGATGTCACTAAATCTTGAGAAAGCTGACTTACTA 827
 Db 61 LysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 80
 QY 828 AAGGCTATTCAAGAACAAATTTGATCGCTAACGTCACAGTAAAGCAGCTACTTTGAGTGC 887
 Db 81 LysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAspLysPheGlnVal 100
 QY 888 ATTGATTTTGCAGGATGCAACCATTAAGTCAAGGAGGAGGCTACTTTGCTGATC 947
 Db 101 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValValTyrPheAlaAsp 120
 QY 948 AAAGATGGTTCGGTAACTTGGCCGACCCCACTGTGCCAAGATTTTGTACCGGACAT 1007
 Db 121 LysAspGlySerValThrLeuProThrGlnProValGlnGlnPheLeuLeuSerGlyHis 140
 QY 1008 GTGGCGCTTAGACCATATAAAGAAACCAATTAACAACCAAGCGGAATCTGTGATGTC 1067
 Db 141 ValAlaGlnAlaArgProTyrLysGlnLysProIleGlnAsnGlnAlaLysSerValAspVal 160
 QY 1068 GAATATCTGTGACGTTTACTCCCTTAAACCTGATGAGATTCGATTCAGACGGTCAAA 1127
 Db 161 GlnTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 180
 QY 1128 GATTAAGTAACTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGATTAATA 1187
 Db 181 AspThrLysLeuLeuThrLeuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 200
 QY 1188 GCTGACGACCAAGCATTTTAAACAACCAACCCAGGCTATACGATTTATGACGTCAC 1247
 Db 201 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnLysAsp 220
 QY 1248 TCCTCAATCGTCACTCATGATGACATGATCTTCGATGAGATTTTACCATGATGATCAAGG 1307
 Db 221 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 240

QY 1308 TTACTTACCGTGTAAAAATCGGGAACAGCTTATAGATCAATAAAAATCTGTCG 1367
 Db 241 PheThrTyrHisValValysAsnArgGlnGlnAlaTyrGlnIleAsnLysSerGlyLeu 260
 QY 1368 AATGAGAAGATTAACAACACTGACCTGATCTCTGAGAAATATTACGCTTTAAAAAGAG 1427
 Db 261 AsnGlnGlnIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 280
 QY 1428 GAAAAGCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACATATAGCTT 1487
 Db 281 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 300
 QY 1488 GATGTCGATACCAACGAATTCCTAAAAAGTAGACGCTTAAACGCTAGGCAACGTAC 1547
 Db 301 AspAlaSerThrAsnGlnLeuLeuLysSerGlnGlnLeuThrAlaSerGlnLysAsn 320
 QY 1548 TTGACTTCAAGATTTATATACGATCTCTGTATAGGCTAAAGCTTACTTACAAATCTC 1607
 Db 321 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 340
 QY 1608 GATGCTTTGCTATTATGACTATACCTTAACTGAAAAATGAGAGATTAATCAGATGAC 1667
 Db 341 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyValGlnAspAsnHisAspAsp 360
 QY 1668 ACCAAGCGTATCAATACCGCTTATATAGGCGAAGGACCCGGAAGAGAAATGCTAGTAC 1727
 Db 361 ThrAsnArgIleIleThrValTyrMetGlyLysArgProGlnGlyGlnAlaSerTyr 380
 QY 1728 CATTTAGCT 1736
 Db 381 HisLeuAla 383

RESULT 3

Q6UK57_STREQ
 ID Q6UK57_STREQ PRELIMINARY; PRT; 436 AA.

AC Q6UK57;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Streptokinase.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35666;
 RA Costa C.S., Torres F.A.G., Filho S.A.;
 RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY368335; AAC73571.1; -; Genomic_DNA.
 DR SMR: Q6UK57; 59-394.
 DR GO: GO:0016301; F:Kinase activity; IEA.
 DR GO: GO:0008243; F:Plasminogen activator activity; IEA.
 DR GO: GO:0005515; F:Protein binding; IEA.
 DR InterPro: IPR004093; Streptokinase.
 DR InterPro: IPR008124; Streptokinase.
 DR Pfam: PF02821; Streptokinase; 3.
 DR PRINTS: PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 436 AA; 49658 MW; CA2DD95F957D8F3C CRC64;

Alignment Scores:
 Pred. No.: 2,346-129 Length: 436
 Score: 1943.00 Matches: 374
 Percent Similarity: 98.43% Conservative: 3
 Best Local Similarity: 97.65% Mismatches: 6
 Query Match: 52.15% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q6UK57_STREQ (1-436)

QY 588 ATTGCTGAGCTGAGTGGCTGAGACCGTCCATCTGTCAACAACGCAATGTTGTT 647

```

Db      23  IleaIaGIProGIuTrIpleuLeuSeArProSeValaSnAnSeRGIleuValI 42
Qy      648  AGCGTTGCGTACTGTTGAGGGGAGCAATCAAGATTAAGCTTTAAATTTTGAATC 707
Db      43  SerValaIaGIYThValGIuGIYThraSnGIaSPILSeRleuYSPhePheGIuIle 62
Qy      708  GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db      63  AspleuThrSeRGIProIaHISGIYGIYSPThGIuGIuIleuSeRProIYSeR 82
Qy      768  AAACCATTTGCTACTGATAGTGGCGGATGTCACATTAACCTTGAGAAAGCTGA 827
Db      83  LysProPheAlaThraSPSeRGIYalameTProHISleYSeuGIuYsaIaSPleu 102
Qy      828  AAGGTATTCAGAACATTAATGCTGAACGTCACAGTACAGACATCTTGAGGTC 887
Db      103  LysaIaIleGIuGIuGIuIleuIleAlaSnValHISeRAsnaSPaSPTYrPheGIuVal 122
Qy      888  ATTGATTTTGAAGCGATCAACCATTAATCGAAGCGGCAAGGCTTACTTGCTGAC 947
Db      123  IleSPheAlaSeRAsnaSPaIaThrIleThraSPaGsaNGIYUValTYrPheAlaSP 142
Qy      948  AAAGATGCTTCGGTAACTTGCCGACCAACCTGTCGAAGATTTTGTCAAGCGGACAT 1007
Db      143  LysaSPGIYSeRValThreuProThrGIuProValGIuGIuPheleuSeuYSGIYHIS 162
Qy      1008  GTGCCCGTTAAGCACTATTAAGAAAAACCAATACAAACCAAGCAATCTGTTAGTGG 1067
Db      163  ValaIraValaIraSProTYrIleSGIuYSProValGIuSaNGIuIaIaYSeRValaSPVal 182
Qy      1068  GAATATACGTGTAAGTTACTCCCTTAACCCGATGAGATTTGAGACAGGCTTCAA 1127
Db      183  GIuTYrThraValGIuPheThraProleuSaSPaSPaSPaSPaSPaSPaSPaSPaSPa 202
Qy      1128  GATACTAAGCTAATGAAAAACATAGCTATCGGTGACACACATCAATCTTAAGATTA 1187
Db      203  AsPThrIYleuIleuYSeRThreuAlaIleGIYAsPThrIleThSeRGIuGIuIleu 222
Qy      1188  GCTCAAGCAAAAGCATTTTAAACAAACCAACCCGCTATACGATTTATGACGCTGAC 1247
Db      223  AlaGIuIaGIuIleuSeRleuSeuYSPThraSPaSPaSPaSPaSPaSPaSPaSPaSPa 242
Qy      1248  TCCCTAATCGTACCTGACATGACATGATTTCCGTAAGATTTTACATGATGATCAAG 1307
Db      243  SeRSeRleIeValThraSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPa 262
Qy      1308  TTTACTTACCGTGTAAATAATCGGAGACAAAGCTTAAGATCAATTAATAATCGGTC 1367
Db      263  PheThrIYThraValaIleuSaNGIuGIuIaIaIaIaIaIaIaIaIaIaIaIaIa 282
Qy      1368  AATGAAGAATTAACACATGACCTGATCTGTGAGAAATATTACTCTTAAAAAGG 1427
Db      283  AsnGIuGIuIleuSaNGIuIleuSaNGIuIleuSaNGIuIleuSaNGIuIleuSaNGI 302
Qy      1428  GAAAGCCGCTATGATCCCTTGAATGGCAGTCACTTGAACCTGTCACCATCAATCG 1487
Db      303  GIuYSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSP 322
Qy      1488  GATGTCGATTAACCAAGATTCCTTAAAGAGTGAAGCTCTTAACAGCTTACGCAAG 1547
Db      323  AsPValaIaThraSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPa 342
Qy      1548  TTAGACTTCAGAGATTTATACGATCTCGTGAATAAGGCTTAACTACTTACAACAAT 1607
Db      343  LeuSPheThraSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPa 362
Qy      1608  GATGCTTTTGGTATTAAGCATATACCTTAATCGAAAGATGAGATTAATCAAGTAC 1667
Db      363  AsPAlaIaPheGIYIleMeCaSPTYrThraSPaSPaSPaSPaSPaSPaSPaSPaSPa 382
Qy      1668  ACCAACCGTATCATTAACGTTATATGAGCAAGCGACCGAAGAGAGATGCTACTAC 1727

```

```

Db      383  ThraSPaNGIleIleThraValTYrMeCGLYsaSPaSPaSPaSPaSPaSPaSPa 402
Qy      1728  CATTAGCT 1736
Db      403  HisleuAla 405

RESULT 4
STRP_STRS1
ID STRP_STRS1 STANDARD; PRT: 440 AA.
AC P10519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Streptokinase G precursor.
GN Name=skg;
OS Streptococcus sp. (strain 19909).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=69017;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89160265; PubMed=2922269;
RA Walter F., Siegel M., Malke H.;
RT "Nucleotide sequence of the streptokinase gene from a group-G
  Streptococcus.";
RL Nucleic Acids Res. 17:1262-1262(1989).
CC -!- FUNCTION: This protein is not a protease, but it activates
  plasminogen by complexing with it. As a potential virulence
  factor, it is thought to prevent the formation of effective fibrin
  barriers around the site of infection, thereby contributing to the
  invasiveness of the cells.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use as long as its content is in no way modified and this statement is not
  removed.
CC -----
DR EMBL; X13400; CAJ31766.1; -; Genomic DNA.
DR PIR; S02723; S02723.
DR HSSP; P00779; 1QOR.
DR SMR; P10519; 63-398.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KM Plasminogen activation; Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase G.
SQ SEQUENCE 440 AA; 50199 MW; 5521F8825FE1B6EA CRC64;

Alignment Scores:
Pred. No.: 2,76e-129 Length: 440
Score: 1942.00 Matches: 375
Percent Similarity: 98.43% Conservative: 2
Best Local Similarity: 97.91% Mismatches: 6
Query Match: 52.12% Indels: 0
Gaps: 0

US-09-940-235-12 (1-2096) x STRP_STRS1 (1-440)
Qy      588  ATTGCTGAGCTGAGTGGCTGCTGACCGCTCATCTGTCAACAAAGCAATGGTGTGTT 647
Db      27  IleaIaGIProGIuTrIpleuLeuSeArProSeValaSnAnSeRGIleuValI 46
Qy      648  AGCGTTGCTGTACTGTTGAGGGGAGCAATCAAGATTAAGCTTTTGAATC 707
Db      47  SerValaIaGIYThraValGIuGIYThraSnGIaSPILSeRleuYSPhePheGIuIle 66
Qy      708  GATCTAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db      67  AspleuThrSeRGIProIaHISGIYGIYSPThGIuGIuIleuSeRProIYSeR 86

```

```

QY 768 AAACATTGCTACTGATAGTGGCGGAGTGCATTAACCTTGAGAAAGCTGACTACTA 827
Db 87 LysleuPhealathraspergilyAlameProhibLysleuGluysAlaAspleuLeu 106
QY 828 AAGGCTATTCAAGAAATTTGATGCGTAAAGTCAAGCTAAAGCACTTTGAGGTC 887
Db 107 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspArgIyrPheGluVal 126
QY 888 ATTGATTTTGAAGCGATGCAACCTTACTGATGAAACGGGAGGTCCTTTGCTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIyrPheAlaAsp 146
QY 948 AAAGATGTTGCGTAACTTTGCGGACCCAACTGTCGAGAAATTTTGTAAAGGACAT 1007
Db 147 LysAspGlySerValThrleuProIleGlnProValGlnGlnPheLeuLeuLysGlyHis 166
QY 1008 GTGGCGGTTAGACCATATAAGAAAAACAATACAAACCAAGCGAAATCTTTGATGTG 1067
Db 167 ValArgValArgProIyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1068 GAATATAGTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGAGTGTCAA 1127
Db 187 GluIyrThrValGlnPheThrProLeuAsnProAspAspPheArgProAlaLeuLys 206
QY 1128 GATACTAAGCTATTGAAAACTAGCTATCGGTGACACATCATCTCAAGAAATTACTA 1187
Db 207 AspThrLysleuLeuLysThrleuAlaIleGlyAspThrIleThrSerGlnGlnLeuLeu 226
QY 1188 GGTCAAGCAAGCAATTTTAAACAAACCAACCCAGGCTATACGATTATGAAAGCTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyIyrThrIleIyrGlnLysArgAsp 246
QY 1248 TCCCTAATGTCACATGACATGACATGATCTTCCGTAAGCTTTTACCAAGATCAAGAG 1307
Db 247 SerSerIleValIleThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
QY 1308 TTTACTTACCGTGTAAATATCGGGAACAAGCTTATAGATCAATAAATATCGGTCTG 1367
Db 267 PheThrIyrHisValIleAsnAsnThrAspleuIleSerGlnLysIyrValIleuLysGly 286
QY 1368 AATGAAAGAAATTAACAACACTGACTGATCTGTGAGAAATATTACGTCCTTAAAGAGG 1427
Db 287 AsnGlnGlnIleAsnAsnThrAspleuIleSerGlnLysIyrValIleuLysGly 306
QY 1428 GAAAGCCGATATGATCCCTTGTATGCGGATCACTTGAACCTGTTCAACCAATACGT 1487
Db 307 GluLysProIyrAspProPheAspAspSerHisleuLysleuPheThrIleLysIyrVal 326
QY 1488 GATGTCGATACCAAGCAATTTGCTAAAGAGGACGCTTTACAGCTAGCGAAGCTAAC 1547
Db 327 AspValAsnThrAsnGlnLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysArgAsn 346
QY 1548 TTAGACTTCAGAGATTTATATAGTCTGATAGAGGCTAAACTACTCTCAACAATCTC 1607
Db 347 LeuAspPheArgAspleuLysAspProArgAspLysAlaLysleuLeuLysArgAsnLeu 366
QY 1608 GATGCTTTTGTATATGACATTAACCTTAACCTGAGAAAGTAGAGATAATCAAGATAC 1667
Db 367 AspAlaPheGlyIleMetAspIyrThrIleuThrGlyLysValGlnAspAsnHisAspAsp 386
QY 1668 ACCAACCGTATACATACCGTTTATATAGGCAAGCGAACCGGAGAGAGAAATGCTACTAC 1727
Db 387 ThrAsnArgIleIleThrValIyrMetGlyLysArgProGlnGlyGlnAsnAlaSerIyr 406
QY 1728 CATTTAGCT 1736
Db 407 HisLeuAla 409

```

RESULT 5
Q7X0Y1_STRPY
Q7X0Y1_STRPY PRELIMINARY; PRT; 440 AA.
AC Q7X0Y1_01-OCT-2003 (Tremblrel. 25, Created)

```

DT 01-OCT-2003 (Tremblrel. 25, last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, last annotation update)
DE Streptokinase.
CN Name=aka:
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OK NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D488;
RA PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RX Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234137; AAP3957.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SNR; Q7X0Y1; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaing activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004053; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KM Kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4BEB9647043BAC CRC64;

Alignment Scores:
Pred. No.: 3,43e-120 Length: 440
Score: 1814.00 Matches: 347
Percent Similarity: 95.04% Conservative: 17
Best Local Similarity: 90.60% Mismatch: 19
Query Match: 48.68% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q7X0Y1_STRPY (1-440)
QY 588 ATTGCTGACCTGATAGTGGCTGTAGACCGTTCATCTGTCAACAAAGCAATGTTGTT 647
Db 27 IleAlaGlyIyrGlyIyrPleuProAspArgProIleAsnAsnSerGlnLeuVal 46
QY 648 AGCGTGTCTGTACTGTGTTGAGGGAAGCAATCAAGACATTAAGCTTTAAATTTTGAATC 707
Db 47 SerMetAlaGlyIleValGlnGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
QY 708 GATCTAACATACAGACCTGCTCATGAGGAAAGACAGAGAGGCTTAAGTCCAAATCA 767
Db 67 AspleuThrSerGlnProAlaHisGlyLysThrGlnGlnLysSerProLysSer 86
QY 768 AAACATTGCTACTGATAGTGGCGGATGTCATTAACCTTGAGAAAGCTGACTACTA 827
Db 87 LysPhePheAlaThrAspAsnGlyAlaMetProhibLysleuGluysAlaAspleuLeu 106
QY 828 AAGGCTATTCAAGAAATTTGATGCGTAAAGTCAAGCTAAAGCACTTTGAGGTC 887
Db 107 LysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspArgIyrPheGluVal 126
QY 888 ATTGATTTTGAAGCGATGCAACCTTACTGATGAAACGGGAGGTCCTTTGCTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValIyrPheAlaAsp 146
QY 948 AAAGATGTTGCGTAACTTTGCGGACCCAACTGTCGAGAAATTTTGTAAAGGACAT 1007
Db 147 LysAspGlySerValThrleuProThrGlnProValGlnGlnPheLeuLeuLysGlyHis 166
QY 1008 GTGGCGGTTAGACCATATAAGAAAAACAATACAAACCAAGCGAAATCTTTGATGTG 1067
Db 167 ValArgValArgProIyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1068 GAATATAGTGTACAGTTTACTCCCTTAAACCTGATGACGATTTGACAGCGAGTGTCAA 1127

```

Db 187 LysTyrThrValGlnPheThrProLeuAnPProAspAspPheArgProGlyLeuLys 206
 QY 1128 GATACTAAGCTATTGAAAACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1187
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGGCTATTCGATTTATGAAAGCTGAC 1247
 Db 227 AlaGlnAlaGlnSerIleLeuAnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 246
 QY 1248 TCCCTCAATCGTCATGACATGACATGATTTTCCGATCGATTTTACCAATGATGACAG 1307
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 1308 TTTACTTACCGTGTAAATAATCGGGAAACAAGCTTTATAGGATCAATTAATAATCGCTG 1367
 Db 267 PheThrTyrArgValLysAspArgGlnGlnAlaTyrGlyIleAnLysLysSerGlyLeu 286
 QY 1368 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGG 1427
 Db 287 AsnGlnGluIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrIleLeuLysGly 306
 QY 1428 GAAAGCCGTATGATCCCTTTGATGCGCATCTGAAACTGTTCCATCAATACGTT 1487
 Db 307 GluSerProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyrVal 326
 QY 1488 GATGTCGATACCAAGCAATTCGTTAAAGTGACGCTCTTAAACAGCTACGCAAGCTAAC 1547
 Db 327 AspValAsnThrAsnGlnLeuLysSerGlnGlnLeuThrAlaSerGlnAlaGln 346
 QY 1548 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTAAACTACTTACCAACATCTC 1607
 Db 347 LeuAspPheArgAspLeuTyrAspProAspLysValAlaLysLeuLysThrAsnLeu 366
 QY 1608 GATGCTTTGGTATTATGACTATACCTTAACGAAAAGTAAAGATATACGATGAC 1667
 Db 367 AspAlaPheAspIleMetAspTyrThrLeuThrGlyValGlnAspAsnHisAspLys 386
 QY 1668 ACCAAGCTATCATACCGTTTATATGCGGCAACGACCCGAAAGAGAGAATGTTAGCTAC 1727
 Db 387 AsnAsnAlaGlnLeuAlaThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
 QY 1728 CATTTAGCT 1736
 Db 407 HisLeuAla 409
 RESULT 6
 Q7X0Y8_STRPY PRELIMINARY; PRT; 440 AA.
 ID Q7X0Y8_STRPY PRELIMINARY; PRT; 440 AA.
 AC Q7X0Y8;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=eka;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=88-019;
 RA PubMed=14679231; DOI=10.1128/DB.186.1.110-121.2004;
 RA Kalita A.; Bessen D.E.;
 RT "Natural selection and evolution of streptococcal virulence genes
 RT involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL; AY234129; AAP39949.1; -; Genomic_DNA.
 DR HSSP; Q53284; 1CAP.
 DR SMK; Q7X0Y8; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasminogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.

DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KW KINASE.
 SQ SEQUENCE 440 AA; 49885 MW; 69DB44F026E3975 CRC64;
 Alignment Scores:
 Pred. No.: 2,44e-119 Length: 440
 Score: 1802.00 Matches: 347
 Percent Similarity: 94.52% Conservative: 15
 Best Local Similarity: 90.60% Mismatches: 21
 Query Match: 48.36% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x Q7X0Y8_STRPY (1-440)
 QY 588 ATTGCTGACCTTGAAGTGGCTGTAGACCGTTCATCTGTCAACAACAGCAATTGGTGT 647
 Db 27 IleAlaGlyTyrGlyTyrPheLeuProAspArgProValAsnSerGlnLeuVal 46
 QY 648 AGCGTTCGTGTACTGTGGGGGACGAATCAAGACTTGTGCTTAAATTTTGAATC 707
 Db 47 SerMetAlaGlyIleValGlnGlyTyrAspLysValPheIleAsnPheGluIle 66
 QY 708 GATCTTAACATCAACGACCTGCTCATGAGGAGAAAGACAGACGAGGCTTAAGTCAAAATCA 767
 Db 67 AspLeuThrSerGlnHisAlaHisGlyLysThrGlnGlnLeuSerProLysSer 86
 QY 768 AAACATTTGCTATCTGATAGTGGCGGATGTCATATAACTTGAAGAGCTGACTTA 827
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlnLysValAspLeu 106
 QY 828 AAGGCTATTGAAGAATGATGCTGATGCTCAAGTACAGTACGACACTTGTAGAGTC 887
 Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
 QY 888 ATTGATTTTGCAGCGGATGCAACATTCATGATGAAAGGCAAGCTTCTTGGCTGAC 947
 Db 127 IleAspPheAlaSerAspAlaThrIleHisAspArgAsnLysValTyrPheAlaAsp 146
 QY 948 AAGATGCTTCGCTTAACCTTGCACCAACCTGCTCCAGAAATTTTGTAAAGCGGACAT 1007
 Db 147 LysAspGlySerValThrLeuProthrinProValGlnIleLeuLeuSerGlyHis 166
 QY 1008 GTCCGGCTTAGACCATTAAGAAAGAAACCAATACAAACCAAGCAATCTGTGATGTC 1067
 Db 167 ValArgValArgProTyrLysGlnLysProValGlnAsnGlnAlaLysSerValAspVal 186
 QY 1068 GAATATACGTACAGTTTACCTGCTTAACCCGTATGACGATTTTCAGACGAGCTGAAA 1127
 Db 187 LysTyrThrValGlnPheThrProLeuAnPProAspAspPheArgProGlyLeuLys 206
 QY 1128 GATACTAAGCTATTGAAAACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 1187
 Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu 226
 QY 1188 GCTCAAGCAAAAGCATTTTAAACAAAACCAACCGGCTATTCGATTTATGAAAGCTGAC 1247
 Db 227 AlaGlnAlaGlnSerIleLeuAnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 246
 QY 1248 TCCCTCAATCGTCATGACATGACATGATTTTCCGATCGATTTTACCAATGATGACAG 1307
 Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
 QY 1308 TTTACTTACCGTGTAAATAATCGGGAAACAAGCTTTATAGGATCAATTAATAATCGCTG 1367
 Db 267 PheThrTyrArgValLysAspArgGlnGlnAlaTyrGlyIleAnLysLysSerGlyLeu 286
 QY 1368 AATGAAGAATAAACAACATGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGG 1427
 Db 287 AsnGlnGluIleAsnAsnThrAspLeuIleSerGlnLysTyrTyrIleLeuLysGly 306

QY 1428 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTGTAAGTCTTACCACCAATACGTT 1487
 |||
 Db 307 GUSerProtyrAspProPheAspArgSerHisLeuIleuPheThrIleIysTyrVal 326
 |||
 QY 1488 GATGTCGATACCAACGAATTGCTAAAGAGGACAGCTCTTAACAGCTACGGAACGTAAC 1547
 |||
 Db 327 AspValAsnThrAsnGlnIleuLeuIysSerGlnGlnIleuLeuThrAlaSerGlnIleuGln 346
 |||
 QY 1548 TTAGACTTCGAGATTTATATACGATCCCTCGATAGAGCTTAACCTACTCTCAACAATCTC 1607
 |||
 Db 347 LeuAspPheArgAspLeuTyrAspProCysAspIleAlaIleuLeuTyrAsnAsnIleu 366
 |||
 QY 1608 GATGCTTTTGTGATTATGACATATACCTTAACCTGAGAAAGTAGAGATTAATCAAGATGAC 1667
 |||
 Db 367 AspAlaPheAspIleuMetAspTyrThrLeuThrGlyIleValGlnAspAsnHisAspIle 386
 |||
 QY 1668 ACCAACCGTATCATTAACCGTTTATATAGGCAACGACCGCAAGAGAGAAATGCTGTTAC 1727
 |||
 Db 387 AsnAsnArgIleValAlaThrValIleuMetGlyIysArgProIleGlyAlaIysGlySerTyr 406
 |||
 QY 1728 CATTTAGCT 1736
 |||
 Db 407 HisLeuAla 409
 |||
 RESULT 7
 Q8KSR8 STRP3 PRELIMINARY; PRT: 440 AA.
 ID Q8KSR8 STRP3 079W73;
 AC 08KSR8 079W73;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Streptokinase A.
 GN Name=ska; Ordered locus names=SPa1700, SpvM3_1698;
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301448;
 RX NCBI_TaxID=301448;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRPAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
 RA Beres S.B., Glyva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 RL [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRPAIN=SSI-1 / Serotype M3;
 RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
 RA Nakagawa I., Kurikawa K., Yamashita A., Nakata M., Tomiyasu Y.,
 RA Okahashi N., Kawabata S., Yamazaki K., Shibata T., Yasunaga T.,
 RA Hayashi H., Hattori M., Hamada S.;
 RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
 RT large-scale genomic rearrangement in invasive strains and new insights
 RT into phage evolution.";
 RL Genome Res. 13:1042-1055(2003).
 DR EMBL; AE014169; AAM80305.1; -; Genomic_DNA.
 DR EMBL; BA000034; BAC64795.1; -; Genomic_DNA.
 DR HSSP; P00779; 1L4D.
 DR SMR; Q8KSR8; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasmaogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro; IPR004093; Staphylokinase.
 DR InterPro; IPR008124; Streptokinase.
 DR Pfam; PF02821; Staphylokinase; 3.
 DR PRINTS; PR01753; STREPKINASE.
 KM Complete proteome; kinase.
 SQ SEQUENCE 440 AA; 49911 MW; 7CCB44F4026B3975 CRC64;

Alignment Scores:
 Pred. No.: 2,44e-119 Length: 440
 Score: 1802.00 Matches: 347
 Percent Similarity: 94.52% Conservative: 15
 Best Local Similarity: 90.60% Mismatches: 21
 Query Match: 48.36% Indels: 0
 DB: 2 Gaps: 0
 US-09-940-235-12 (1-2096) x Q8KSR8_STRP3 (1-440)
 QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGCTCATCTGTGCAACAAGCCAAATTGGTTGTT 647
 |||
 Db 27 TLeaIaGlyTyrGlyTyrPleuProAspArgProProValAsnAsnSerGlnLeuVal 46
 |||
 QY 648 AGCTTCGTGCTACTGTTGGAGGGACCAATCAAGACATTAGCTTAAATTTTGAATC 707
 |||
 Db 47 SerMetIaGlyIleValIleGlnGlyThrAspIleIleValIleAsnPheGlnIle 66
 |||
 QY 708 GATTAACATCAAGACCGCTCATGAGAGAGACAGAGAGGCTTAAGTCAAAATCA 767
 |||
 Db 67 AspLeuThrSerGlnHisAlaHisGlyGlyIleThrGlnGlnGlnIleuSerProIysSer 86
 |||
 QY 768 AAACCATTTGCTACTGATAGTGGCGCATGTCACATTAACCTTGAAGAGCTGACTCTA 827
 |||
 Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisIleuGlnIleuValAspLeu 106
 |||
 QY 828 AAGCTATTCAGACCAATTGATGCTTACGTCACAGTAACAAGCACTGTTGAGGTC 887
 |||
 Db 107 LysAlaIleGlnIleuGlnIleuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
 |||
 QY 888 ATTGATTTTGCAAGCGATGACCAATTAATGATGGAACCGGACAGCTTACTTGTCTGAC 947
 |||
 Db 127 IleAspPheAlaSerPalaThrIleThrAspArgAsnGlyIleValTyrPheAlaAsp 146
 |||
 QY 948 AAAGATGTTCCGGTAACTTGCACCGACCAACCTGTCCAAAGATTTTGTGAAGGACAT 1007
 |||
 Db 147 LysAspGlySerValAlaThrLeuProThrGlnProValGlnGlnPheLeuSerGlyHis 166
 |||
 QY 1008 GTGGCGGTGTAACCATTAATAAGAAACCAATACAAACCAAGCAAAATCTGTGATGTG 1067
 |||
 Db 167 ValArgValArgProTyrIleuGlnIleuProValGlnAsnGlnAlaIysSerValAspVal 186
 |||
 QY 1068 GAATPACTGTACAGTTTACTCTCTTAAACCTGATGACGATTTCAGACCAAGTCTCAA 1127
 |||
 Db 187 LysIleTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuIys 206
 |||
 QY 1128 GATATTAAGCTATTGAAAACACATAGCTATGGTGACACCAATCATATCTCAAGATTACTA 1187
 |||
 Db 207 AspThrLysLeuLeuIleuIleuValAlaIleGlyAspThrIleThrSerGlnIleuLeu 226
 |||
 QY 1188 GCTCAAGCACAAAGCATTTTAAACAAACCAACCAAGCTTATAGATTATGAAGTGCAC 1247
 |||
 Db 227 AlaGlnAlaGlnIleuThrIleuAsnIleuThrHisIleuProGlyTyrThrIleTyrGlnIleuGln 246
 |||
 QY 1248 TCTCAATGCTCACTCATGACCAATGACATTTTCCGATGCAATTTTACCAATGATCAAGAC 1307
 |||
 Db 247 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGln 266
 |||
 QY 1308 TTACTTACCGGTGTAAATAATCGGGAACAGAGCTTATAGATCAATTAATAATCGTGTG 1367
 |||
 Db 267 PheThrTyrThrValIleuAspArgGlnGlnAlaIleuTyrGlyIleAsnIleuIleuSerGlyLeu 286
 |||
 QY 1368 AATGAAGAATAAACAACATGACCTGATGCTGAGAAATATTCGTCCTTAAAGAGG 1427
 |||
 Db 287 AsnGlnGlnIleuAsnAsnThrAspLeuIleSerGlyIleuTyrTyrIleuIleuIleuGly 306
 |||
 QY 1428 GAAAAGCCGTATGATCCCTTTGATCGAGTCACTGTAAGTCTTACCACCAATACGTT 1487
 |||
 Db 307 GUSerProtyrAspProPheAspArgSerHisLeuIleuPheThrIleIysTyrVal 326
 |||
 QY 1488 GATGTCGATACCAACGAATTGCTAAAGAGGACAGCTCTTAACAGCTACGGAACGTAAC 1547
 |||

Db 327 AspValaenThrAsnGluLeuLeuSerGluGlnLeuLeuThrAlaSerGluArgAsn 346
QY 1548 TTAGACTTCAGAGATTATATACGATCTCGTATAGAGCTAAAGCTACTACCAACATCTC 1607
Db 347 LeuAspPheArgAspLeuThrAspProCysAspLysAlaLeuLeuLeuThrAsnLeu 366
QY 1608 GATGCTTTGGTATATATGACTATACCTTAACCTGGAAGAGATAGATATACAGATGAC 1667
Db 367 AspAlaPheAspPheLeuMetAspPheThrLeuThrGlyValGlnAspAsnHisAspLys 386
QY 1668 ACCAAGCTATATACAGCTTTATATGGCAAGCAGCCAGAGAGAGAAATGCTACTAC 1727
Db 387 AsnAsnArgIleValThrValThrValMetGlyLysArgProLysGlyAlaLysGlySerTyr 406
QY 1728 CATTTAGCT 1736
Db 407 HisLeuAla 409
RESULT 8
STRP_STRPY STANDARD; PRT; 440 AA.
AC P10520;
DT 01-JUL-1989 (Rel. 11, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Streptokinase A precursor.
GN Name=SkA; OrderedLocNames=SPY1979;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SF130/13 / Serotype M1;
RX MEDLINE=89160264; PubMed=2646590;
RA Walter F., Siegel M., Malke H.;
RT "Nucleotide sequence of the streptokinase gene from a Streptococcus
RT Nucleic Acids Res. 17:1261-1261(1989).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21196884; PubMed=11296296; DOI=10.1073/pnas.07159398;
RA Ferreretti J.J., Meshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeux C., Setate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -!- FUNCTION: This protein is not a protease, but it activates
CC plasminogen by complexing with it. As a potential virulence
CC factor, it is thought to prevent the formation of effective fibrin
CC barriers around the site of infection, thereby contributing to the
CC invasiveness of the cells.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL: X13399, CA31765.1, -, Genomic DNA.
CC EMBL: AB006620, AA034665.1, -, Genomic DNA.
CC PIR: S02724, S02724.
CC HSSP: Q53284, 1CAP.
CC SMR: P10520, 63-398.
CC InterPro: IPR004093, Streptolysine.
CC InterPro: IPR008124, Streptokinase.
CC Pfam: PF02821, Staphylokinase, 3.
CC PRINTS: PR01753, STREPKINASE.
CC Complete proteome; Plasminogen activation; Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase A.

FT CONFLICT 163 163 L -> V (in Ref. 1).
FT CONFLICT 345 345 R -> G (in Ref. 1).
FT CONFLICT 373 373 D -> N (in Ref. 1).
FT CONFLICT 428 428 D -> Y (in Ref. 1).
FT CONFLICT 438 438 K -> N (in Ref. 1).
SQ SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;
Alignment Scores:
Pred. No.: 9, 05e-119 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.99% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 48.15% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-12 (1-2096) x STRP_STRPY (1-440)
QY 588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTCACAAACAGCCAAATGGTGT 647
Db 27 IleAlaGlyTyrGlyThrLeuProAspArgProProIleAsnAsnSerGlnLeuVal 46
QY 648 AGCGTTGCTGTACTGTTGAGGAGCAGATCAAGACATTAGTCTTAATTTTGAATC 707
Db 47 SerMetAlaGlyIleValGlnGlyThrAspLysLysValPheIleAsnPhePheGluIle 66
QY 708 GATCTAACATTCACGACCTGCTCAATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 767
Db 67 AspLeuThrSerIleProAlaHisGlyGlyLysThrIleGlnGlyLeuSerProLysSer 86
QY 768 AAACATTTGCTACTATGATGGCGGAGATGTCATTAATCTGAGAAAGCTGACTTCTA 827
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlyAlaAspLeu 106
QY 828 AAGGCTATTCAAGAACAAATGATGCTAAGCTCAAGACGATCTTGAAGTCT 887
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGluVal 126
QY 888 ATGATTTTGCAGAGCAGATGCAACCATTCATGATGCAAGCGCAAGCTTCTGCTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 948 AAAGATGGTGGTGAATCTTGACCGGACCGCAACCGTCCAGAAATTTTGTGAAGGACAT 1007
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluLeuLeuLysGlyHis 166
QY 1008 GTGGCGGTGAACCATATTAAGAAACCAATCAAAACCAAGCGAAATGTTGATGTG 1067
Db 167 ValArgValArgProTyrLysGlyLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1068 GAATATACCTGTACAGTTTACTCCCTTAACCCGTATGACATTTGACAGCGCTGAAA 1127
Db 187 GluTyrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1128 GATCTAAGCTATTGAAAACATGAGCTATGGTGAACCATCATCATGCTAAGATTTCTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlyLeu 226
QY 1188 GCTCAAGCAAAAGCATTTTAAACAAACCAACCAAGCTATGACATTTTGAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlnArgAsp 246
QY 1248 TCTCAATGCTCACTATGACAAATGATCTTTCCGATCGATTTTCAATGATCAAGAG 1307
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1308 TTACTTACCGTGTAAATATGGGAAACAGCTTATAGATCAATTAATAATGCGTGTG 1367
Db 267 PheThrThrHisValLysAsnArgGlnGlnAlaTyrGlnLysProLysPheGlyIle 286
QY 1368 AATGAAGAATAAACAACATGATCTGAGAAATATTTAAGTCTTAAAAAAGG 1427
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGlnLysTyrTyrValLeuLysGlnGly 306


```

QY 1428 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTGAAAGCTTGCAACCATCAATACGTT 1487
Db 307 GtulyserProtyrAspProPheAspArgSerHisLeuLeuLeuPheThrIleuTyVal 326
QY 1488 GATGTCGATACCAAGCAATTTGCTAAAGAGGACACTCTTAACACTGACCGCAACGTAAC 1547
Db 327 AspValAsnThrAsnGluLeuLeuLeuSerGluGlnLeuLeuThrIleSerGluArgAsn 346
QY 1548 TTAGACTTCAGAGATTATATACGATCCTCGATTAAGGCTTAACACTCTCAACAATCTC 1607
Db 347 LeuAspPheArgAspLeuLeuTyrrAspProArgAspIleAlaLeuLeuLeuTyrrAsnAsnLeu 366
QY 1608 GATGCTTTGTGATTATGCACTATACCTTAAGGAAAAGTAGAGATAATCAAGATGAC 1667
Db 367 AspAlaPheAspIleMetAspTyrrThrLeuThrGluValGluAspAsnHisAspIle 386
QY 1668 ACCAACCGTATCATTAACGCTTTATATAGGCAAGCGACCGCAAGAGAGAATGCTAGTAC 1727
Db 387 AsnAsnArgValValThrValTyrrMetGluLeuArgProLeuSerGluAlaLeuSerGlySerTyrr 406
QY 1728 CATTTAGCT 1736
Db 407 HisLeuAla 409

RESULT 9
Q532X6_STRPY PRELIMINARY; PRT; 440 AA.
ID 0532X6_STRPY
AC 0532X6_
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86-779;
RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kalita A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234128; AAP3948.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 440 AA; 49924 MW; D6227BF040B758DB CRC64;

Alignment Scores:
Pred. No.: 9.05e-119 Length: 440
Score: 1794.00 Matches: 341
Percent Similarity: 93.99% Conservative: 19
Best Local Similarity: 89.03% Mismatches: 23
Query Match: 48.15% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q532X6_STRPY (1-440)
QY 588 ATTGCTGACCTGATGCTGCTAGACGTCATGCTGCAACAACAGCAATGTTGTT 647
Db 27 IleAlaGlyTyrrGlyTrrPheProAspArgProPheIleAsnHisSerGlnLeuVal 46
QY 648 AGCGTGTGCTACTGTTGAGGGAGCAATCAAGCATTAAGTCTTAATTTTGAATC 707
Db 47 SerMetAlaGlyIleValGluGlyThrAspLeuValPheIleAsnHisPheGluIle 66
QY 708 GATCTAACATCAACGCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATCA 767
Db 67 AspLeuThrSerGlnProIleHisGlyGlySerThrGluGlnGlyLeuSerProLeuSer 86
QY 768 AAACCATTTGCTACGATGATGCGGAGATGTCATTAACCTTGAGAAAGCTGACTTACTA 827

```

```

Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGluValAspLeuLeu 106
QY 828 AAGGCTATTCAAGAACAAATGATGATGATCCAGAGTAACGACGACTTATGAGGTC 887
Db 107 LysAlaIleGlnLysGlnLeuIleAlaAsnValHisSerAsnAspGlyTyrrPheGluVal 126
QY 888 ATTGATTTTTCAGACGATGCAACCATTAATGATGCAAGCGCAAGCTCTTGTCTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyValTyrrPheAlaAsp 146
QY 948 AAAGATGCTTGGTATCACTTGGCCGACCCAACTCTCCAGAAATTTTGTCAAGCGGACAT 1007
Db 147 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 166
QY 1008 GTGCGCGTTAGACATTTAAAGAAACCAATACAAACCAAGGAAATGCTTGATGTCG 1067
Db 167 ValArgValArgProTyrrLysGluLysProValGlnAsnGlnAlaLysSerValAspVal 186
QY 1068 GAATATATCTGATACGTTACTCCTTAACCGCTGATGACGATTTCAAGCAGTCTCAAA 1127
Db 187 GluTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLys 206
QY 1128 GATATCAAGCTATTGAAACATAGCTATGCTGAGACCATCAATCATCTCAAGAAATTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnLeuLeu 226
QY 1188 GCTCAACGACAAAGCATTTTAAACAAACCAACCGGCTATACGATTTATGAAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrrThrIleTyrrGluArgAsp 246
QY 1248 TCCGTAATCGTCACTCATGACAATGACATTTTCCGTCAGATTTTACCAAGGATCAAGAG 1307
Db 247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1308 TTACTTACCGGTGTTAAAAATCGGAAACAAGCTTATGATCAATAAAAAATCGTCTG 1367
Db 267 PheThrTyrrHisValLysAsnArgGlnGlnAlaTyrrGluIleAsnProLysThrGlyIle 286
QY 1368 AATGAGAAATTAACAACATGACCTGATCTCTGAGAAATTTACGTCCTTAAAAAAGG 1427
Db 287 LysGluLysThrAsnAsnThrAspLeuValSerGluLysTyrrValLeuLysGlnGly 306
QY 1428 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTGAAAGCTTGCAACCATCAATACGTT 1487
Db 307 GtulyserProtyrAspProPheAspArgSerHisLeuLeuLeuPheThrIleuTyVal 326
QY 1488 GATGTCGATACCAAGCAATTTGCTAAAGAGGACACTCTTAACACTGACCGCAACGTAAC 1547
Db 327 AspValAsnThrAsnGluLeuLeuLeuSerGluGlnLeuLeuThrIleSerGluArgAsn 346
QY 1548 TTAGACTTCAGAGATTATATACGATCCTCGATTAAGGCTTAACACTCTCAACAATCTC 1607
Db 347 LeuAspPheArgAspLeuLeuTyrrAspProArgAspIleAlaLeuLeuLeuTyrrAsnAsnLeu 366
QY 1608 GATGCTTTGTGATTATGCACTATACCTTAAGGAAAAGTAGAGATAATCAAGATGAC 1667
Db 367 AspAlaPheAspIleMetAspTyrrThrLeuThrGluValGluAspAsnHisAspIle 386
QY 1668 ACCAACCGTATCATTAACGCTTTATATAGGCAAGCGACCGCAAGAGAGAATGCTAGTAC 1727
Db 387 AsnAsnArgValValThrValTyrrMetGluLysArgProLysGluAlaLeuSerGlySerTyrr 406
QY 1728 CATTTAGCT 1736
Db 407 HisLeuAla 409

RESULT 10
Q7X0Y2_STRPY PRELIMINARY; PRT; 440 AA.
ID Q7X0Y2_STRPY
AC Q7X0Y2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=D306;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Besen D.B.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL: AY234136; AAP39956.1; -; Genomic_DNA.
 DR HSSP: O53284; 1CAP.
 DR SMK; Q7X0Y2; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasma:inogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro: IPR004093; Staphylokinase.
 DR Pfam: PF02821; Staphylokinase; 3.
 DR PRINTS: PRO1753; STRBPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50188 MW; 4CB42AC2A9062C2E CRC64;

Alignment Scores:

Pred. No.: 2,84e-118 Length: 440
 Score: 1787.00 Matches: 342
 Percent Similarity: 94.26% Conservative: 19
 Best Local Similarity: 89.30% Mismatches: 22
 Query Match: 47.96% Indels: 0
 DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q7X0Y2_STRPY (1-440)

QY 588 ATGTGACCGTGAAGTGGTGTGACCGTGCATGTCAGCAAGCAATTTGTTT 647
 DB 27 TTTTGGT 46
 QY 648 AGCGT 707
 DB 47 TTTTGGT 66
 QY 708 GATCTACATGACGACCTGCTCATGAGCAAGCAAGCAAGCTTAAATGTC 767
 DB 67 TTTTGGT 86
 QY 768 AAACATTGT 827
 DB 87 TTTTGGT 106
 QY 828 AAGCTATTCAAGCAATTTGATCGTCAAGCTGTCACAGTAAAGTGTGAGTC 887
 DB 107 TTTTGGT 126
 QY 888 ATTGATTTGCAAGGATGCAACCATTAAGTGAAGGAGGAGTCTTGTGAC 947
 DB 127 TTTTGGT 146
 QY 948 AAAAGT 1007
 DB 147 TTTTGGT 166
 QY 1008 GTGAGCGTTAGACCATTAAGCAAGCAAGCAAGCAAGCAAGCAAGTGTG 1067
 DB 167 TTTTGGT 186
 QY 1068 GAATATGCTGTGACGTTTACTCCCTTAAACCTGATGAGATTTGACAGCTCA 1127
 DB 187 TTTTGGT 206

QY 1128 GATACCTAAGCTATTGAAAACACTAGCTATGCGTGACACATCACTCAAGTACTA 1187
 DB 207 TTTTGGT 226
 QY 1188 GCTTGAAGCAAGCAATTTTAAACAAAACACCCAGGCTTACGATTTTGAAGT 1247
 DB 227 TTTTGGT 246
 QY 1248 TCCCTAAGT 1307
 DB 247 TTTTGGT 266
 QY 1308 TTTTGGT 1367
 DB 267 TTTTGGT 286
 QY 1368 AATGAAGAAATAAACAACACTGACCTGATCTGTGAGAAATATTACGCTTAA 1427
 DB 287 TTTTGGT 306
 QY 1428 GAAAACCGGTATGATCCCTTGTATGCGACGTCATTTGAACGATCAATACGT 1487
 DB 307 TTTTGGT 326
 QY 1488 GATGTGATPACCAAGCAATGTCTTAAAGTGAAGGAGGAGCTTAAAGCTAAC 1547
 DB 327 TTTTGGT 346
 QY 1548 TTAGACTTCAAGATTTATACGATCTCGTATAGGCTTAACTTCAACATCTC 1607
 DB 347 TTTTGGT 366
 QY 1608 GATGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1667
 DB 367 TTTTGGT 386
 QY 1668 ACCAACGATATCAACCGTTTATGAGGCAAGCAAGCAAGCAAGTGTAGCT 1727
 DB 387 TTTTGGT 406
 QY 1728 CATTTAGCT 1736
 DB 407 TTTTGGT 426

RESULT 11
 ID Q7X0Y2_STRPY PRELIMINARY; PRT; 440 AA.
 AC Q7X0Y2;
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Streptokinase.
 GN Name=ska;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IRP112;
 RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
 RA Kalia A., Besen D.B.;
 RT "Natural selection and evolution of streptococcal virulence genes
 involved in tissue-specific adaptations.";
 RL J. Bacteriol. 186:110-121(2004).
 DR EMBL: AY234130; AAP39950.1; -; Genomic_DNA.
 DR HSSP: O53284; 1CAP.
 DR SMK; Q7X0Y7; 63-398.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008243; F:plasma:inogen activator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR InterPro: IPR004093; Staphylokinase.
 DR InterPro: IPR008124; Streptokinase.

DR Pfam: PF02821; Staphylokinase; 3.
 DR PRINTS: PR01753; STREPKINASE.
 KW Kinase.
 SQ SEQUENCE 440 AA; 50160 MW; 1CSAF07907EC7AC8 CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 1758.00	3,276-116	440	338	19	26	0
Percent Similarity: 93.21%						
Best Local Similarity: 88.25%						
Query Match: 47.18%						

US-09-940-235-12 (1-2096) x Q7X0Y7_STRPY (1-440)

```

QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCATCTGTCAACAACAGCAATTGTTGTT 647
DB 27 lleaiaGlyTyrGlyTyrpleuLeuAspArgProValAsnAsnSerGlnLeuValVal 46
QY 648 AGCGTGTCTGTACTGTGTAGGGGAGCAATCAAGACATTAGTCTTAATTTTGAATC 707
DB 47 SerMetAlaGlyLeValAlaGlnGlyThrAspLeuValValPheLeuPheGlnIle 66
QY 708 GATCTAACATCAGACCTGCTCATGAGAGAAAGACAGAGAGGCTTAACTCCAAATCA 767
DB 67 AspLeuThrSerGlnProAlaHisGlyGlyLeuThrGlnGlnGlyLeuSerProLysSer 86
QY 768 AAACCACTTTGCTCATGTAGTGGCGGAGATGCTCATTAATTTGAGAAAGCTTACTA 827
DB 87 LysProPheAlaThrAsnSerSerAlaMetProHisLeuLeuLysValAspLeuLeu 106
QY 828 AAGGCTATTCAAGAACATTTGATGCTAAGCTCAGCAGTACAGCAGCTACTTTGAGTGC 887
DB 107 LysAlaAlaGlnGlnGlnArgLeuLeuAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 888 ATTGATTTTTCAGAGCGATGCAACATTACTGATGGAACCGGCAAGGCTTACTTTGCTGAC 947
DB 127 lLeaSPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAsp 146
QY 948 AAAGATGCTTGGCTTAACCTTGGCCAGCCCAACCTGCTCAGAGATTTTTCCTAACCGGACAT 1007
DB 147 LysAspAspSerValThrLeuProThrGlnProValGlnGlnLeuLeuArgGlyHis 166
QY 1008 GTGCGGCTTACAGCATATTAAGAAACCAATGACAAACCAACGGAATCTGTGATGTCG 1067
DB 167 ValAlaGValAlaArgProLysGlyLysProLysGlnThrProAlaLysSerValAlaPVal 186
QY 1068 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAGACCGGCTTCAAA 1127
DB 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProValAlaLeuLys 206
QY 1128 GATCTAACCTATTGAAAACACTAGCTATGCTGAGACCATCATCTCAAGAAATTACTA 1187
DB 207 AsnThrLysLeuLeuLeuThrLeuAlaAlaGlyGlyThrValThrSerGlnGlnLeuLeu 226
QY 1188 GCTCAAGCAACAAGCAATTTTAAACAAAACCAACCGGCTTATCGATTTTGAAGCTGAC 1247
DB 227 AlaGlnAlaGlnSerLysLeuAsnGlnSerHisAspAspArgTyrThrIleTyrGlnLysAsp 246
QY 1248 TCCCTCAATCGTCACTGACATGACATGATTTTCCGTACAGATTTTCAATGATGATCAAGAG 1307
DB 247 SerSerLysValThrHisAspAsnAspLysPheAspArgThrLysLeuProMetAspGlnGln 266
QY 1308 TTACTTACCGTGTAAATAATGGGAAACAAGCTTATAGATCAATTAATAAATCGTCTG 1367
DB 267 PheThrTyrHisIleLeuAspArgGlnGlnAlaTyrGlyTyrLeuLysLysSerGlyGln 286
QY 1368 AATGAAGAAATTAACAACACTGACTGATCTGAGAAATATTAAGTCTTAAAAAGG 1427
DB 287 GlnGlnLysThrAsnAsnThrAspLeuLysSerGlnLysTyrTyrValLeuLysGly 306
QY 1428 GAAAGCGGTATGATCCCTTATGTCAGACCTTGAACCTTGAACCTTGAACCTTGAACCTT 1487

```

```

DB 307 GlnLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrVal 326
QY 1488 GATGTCGATACCAACCAATTTGCTAAAGAGGAGCGCTTAAACGCTTACAGCTACGAACTGAC 1547
DB 327 AspValAsnThrAsnLysLeuLysSerGlnGlnLeuLeuThrAlaSerGlnLysAsn 346
QY 1548 TTAGACTTTCAGAGATTTTATACGATCCCTGATTAAGGCTTAAACTACTTCAACAATCTC 1607
DB 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1608 GATGCTTTTGGTATTATGACTATATCTTACTTACGAAAGTAAAGATTAATCAGCATGAC 1667
DB 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
QY 1668 ACCAACCGTATCATTAACCGTTTATAGGCGAGACCGAAGAGAGGAATGCTAGCTAC 1727
DB 387 AsnAsnArgValAlaThrValTyrMetGlyLysArgProGlnGlnAlaSerTyr 406
QY 1728 CATTTAGCT 1736
DB 407 HisLeuAla 409

RESULT 12
Q5X9T6_STRP6 PRELIMINARY; PRT; 440 AA.
ID Q5X9T6_STRP6
AC Q5X9T6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Streptokinase (EC 3.4.-.-).
GN Ordered locus names=M6_Spy1692;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGA510394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.E.,
RA Voylich J.M., DeLeo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
RT metagenome: complete genome sequence of a macrolide-resistant serotype
RT M6 strain".
RT M6 strain".
RL J. Infect. Dis. 190:727-738(2004).
DR EMBL; CP000003; AAT87827.1; -; Genomic_DNA.
DR SMR; Q5X9T6; 63-398.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR GO; GO:0005151; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW complete proteome; Hydrolase; Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1CSAF07907EC7AC8 CRC64;

Alignment Scores:
Pred. No.: 3,276-116 Length: 440
Score: 1758.00 Matches: 338
Percent Similarity: 93.21% Conservative: 19
Best Local Similarity: 88.25% Mismatches: 26
Query Match: 47.18% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q5X9T6_STRP6 (1-440)
QY 588 ATTGCTGACCTGAGTGGCTGTAGACCGTCATCTGTCAACAACAGCAATTGTTGTT 647
DB 27 lleaiaGlyTyrGlyTyrpleuLeuAspArgProValAsnAsnSerGlnLeuValVal 46
QY 648 AGCGTGTCTGTACTGTGTAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAATC 707

```

```

Db      47 SerMetAlaGlyIleValIleGluGlyThrAspLysLysValPheIleAsnPheNegIuile 66
Qy      708 GATCTAAACATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATATCA 767
Db      67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnIleLysLeuSerProLysSer 86
Qy      768 AAACCATTTGCTACTATAGTGGCGGATGTCACATAACTTGAAGAGCTGACTACTA 827
Db      87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Qy      828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACACGACGACTTCTTGAGGTC 887
Db      107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrrPheGluVal 126
Qy      888 ATTGATTTTGGCAAGCATGACCATTTACTGATCGAAAGCGGAGGCTTACTTGGTCAC 947
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrrPheAlaAsp 146
Qy      948 AAAGATGGTTGGGTAACTTGGCGGACCGGACCGTCGACAAATTTTGGTAAAGCGGACAT 1007
Db      147 LysAspAspSerValThrLeuProHisGlnProValGlnGluPheLeuLeuArgGlyHis 166
Qy      1008 GTGCGCGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 1067
Db      167 ValArgValArgProTyrrLysGluLysProIleGlnThrProAlaLysSerValAspVal 186
Qy      1068 GAATTTACTGTCACGTTTACTCCCTTAAACCTGATGACGATTTCCAGACCGGCTCAAA 1127
Db      187 ArgTyrrThrValGlnPheThrProLeuAsnProAspAspPheArgProValLeuLys 206
Qy      1128 GATCTAAGCTATTGAAAACCATGACTACGCTGACACCATCATCTCTCAAGAAATTAATA 1187
Db      207 AsnThrLysLeuLeuLysThrLeuAlaIleGlyGlyThrValThrSerGlnGluLeuLeu 226
Qy      1188 GCTCAAGACCAAGACATTTTAAACAAACCAACCCAGCGCTATACGATTTATGAAACGTCAC 1247
Db      227 AlaGlnAlaGlnSerIleLeuAsnGlnLysSerHisProAspTyrrThrIleTyrrGluAsp 246
Qy      1248 TCTCTAATGCTACATCAGTACCAATGACATTTTCCGTAAGATTTTACCAATGATCAAGAG 1307
Db      247 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
Qy      1308 TTTACTTACCGCTTAAATAATGGGAAACAGCTTATAGATCAATTAATAATGCTGTCG 1367
Db      267 PheThrTyrrHisIleLysAspArgGlnGlnAlaTyrrGlyIleAsnLysLysSerGlyGln 286
Qy      1368 AATGAAGAAATTAACACACATGACCTGATCTCTGAGAAATTTAAGCTTAAATAAGG 1427
Db      287 GlnGluLysThrAsnAsnThrAspLeuIleSerGluLysTyrrTyrrValLeuLysLysGly 306
Qy      1428 GAAAAACCGTATGATCCCTTTGATGCGACGACTCATGMAAATGCTTCCATCAATACGTT 1487
Db      307 GlnLysProTyrrAspProPheAspArgSerHisLeuLysLeuPheThrIleAsnTyrrVal 326
Qy      1488 GATGTCGATCCAAAGAAATTTGCTAAATAATGAGCAGCTTTCACAGCTTCCGCAAGCTAAC 1547
Db      327 AspValAsnThrAsnLysLeuLysSerGlnGlnIleLeuThrLaseGluAlaGaa 346
Qy      1548 TTAGACTTCAGAGATTTATAGCATCCTGCTGATAAGGCTAAATCTACTACACAAATCTC 1607
Db      347 LeuAspPheArgAspLeuTyrrAspProAspAspLysAlaLysLeuLeuTyrrAsnAsnLeu 366
Qy      1608 GATGCTTTTGTATTATGACTATACCTTACTGAGAAAGTAGAGATATACAGATGAC 1667
Db      367 AspAlaPheGlyIleMetAspTyrrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
Qy      1668 ACCAACCGTATCATTAACGTTTATATGGCAAGCGAACCCGAAAGAGAGATGCTACTAC 1727
Db      387 AsnAsnArgValValThrValTyrrMetGlyLysArgProGlnGlyGluAsnAlaSerTyrr 406
Qy      1728 CATTTAGCT 1736

```

```

Db      407 HisLeuAla 409
RESULT 13
Q8NZAG_STRP8 PRELIMINARY; PRT; 440 AA.
ID Q8NZAG_STRP8 PRELIMINARY; PRT; 440 AA.
AC Q8NZAG;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Streptokinase.
GN OrderedLocustNames=spym18_2042;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=301451;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Forceella S.F.,
RA Parkins L.D., Berez S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE010107; AAL98517.1; -, Genomic_DNA.
DR HSSP; O53284; 1C4P.
DR SMR; Q8NZAG; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Streptokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Streptokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Complete proteome; kinase.
SQ
SEQUENCE 440 AA; 50382 MW; 47DAB35577EBB88E CRC64;

Alignment Scores:
Pred. No.: 5,34e-116 Length: 440
Score: 1755.00 Matches: 334
Percent Similarity: 93.738 Conservative: 25
Best Local Similarity: 87.218 Mismatches: 24
Query Match: 47.108 Indels: 0
DB: 2 Gaps: 0

US-09-940-235-12 (1-2096) x Q8NZAG_STRP8 (1-440)
Qy      588 ATTGCTGACCTGAGTGGCTGCTAGACCGTCACTGTCACAAACAGCAATGTTGTT 647
Db      27 IleAlaGlyTyrrGluThrLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
Qy      648 AGCGTTGCTGCTACTGTTTGAAGGGAAGATCAAGATTAAGTCTTAAATTTTGAAGC 707
Db      47 SerMetAlaGlyIleValIleGluGlyThrAspLysLysValPheIleAsnPheNegIuile 66
Qy      708 GATCTAATCAATCAAGACCTGCTCATGAGAGAAAGACAGACAGGCTTAAGTCCAAATATCA 767
Db      67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnIleLysLeuSerProLysSer 86
Qy      768 AAACCATTTGCTACTATAGTGGCGGATGTCACATAACTTGAAGAGCTGACTACTA 827
Db      87 LysProPheAlaThrAsnSerSerAlaMetProHisLysLeuGluLysAlaAspLeuLeu 106
Qy      828 AAGGCTATTCAAGAACATTTGATCGCTAACGTCACAGTACACGACGACTTCTTGAGGTC 887
Db      107 LysAlaIleGlnGluArgLeuIleAlaAsnValHisSerAsnAspGlyTyrrPheGluVal 126
Qy      888 ATTGATTTTGGCAAGCATGACCATTTACTGATCGAAAGCGGAGGCTTACTTGGTCAC 947
Db      127 IleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrrPheAlaAsp 146

```

```

QY 948 AAGATGTTGCGTAACTTTCGCCGACCACTGTCCAGAAATTTTGTGAAGCGGACAT 1007
Db 147 GlnAsglySerValThrLeuProThrGlnProIleGlnGlnPheLeuAlaGlyHis 166
QY 1008 GTGGCCGTTAGAACATTAAGAAAAACCAATCAAAACCAAGCGAAATCTGTGATGTG 1067
Db 167 ValArgValAlaGlyProTyrGlySgIlySgIlyProIleGlnThrProAlaLysSerValAspIle 186
QY 1068 GAATATACGTGTACAGTTACTCCCTTAACCCGTAGACGATTTCAGACAGGCTGAAA 1127
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY 1128 GATATTAAGCTATTGAAAAACATAGCTATCGGTGACACATCAATCTCAAGATTACTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 1188 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTATGAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisSerAspTyrThrIleTyrGlnArgAsp 246
QY 1248 TCCTCAATGCTACATGACATGACATTTTCGTAAGCTTTCACCAATGATGACAGAG 1307
Db 247 SerSerIleValAlaThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 266
QY 1308 TTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGATCAATAAAAATCTGGCTCG 1367
Db 267 PheThrTyrHisIleLeuAspArgGlnGlnAlaTyrGlyIleAsnLysLysSerGlyGln 286
QY 1368 AATGAAGAAATTAACAACAACATGACCTGATCTGTGAGAAATATTAAGCTTTAAAGAGG 1427
Db 287 GlnGlnLysThrAsnAsnThrAspLeuIleSerGlnLysTyrTyrValLeuLysGly 306
QY 1428 GAAAAGCCGTATGATCCCTTTGATCGCACTGACTTGAACCTGTTCACCATCAATACGTT 1487
Db 307 GlnLysProTyrAspProPheAspArgSerHisLysLeuLysPheThrIleAsnTyrVal 326
QY 1488 GATGTCGATACCAACGAATTTGCTAAAGAGTACAGCTCTTAACAGCTAGCGAAGCAAC 1547
Db 327 AspValAsnThrAsnLysLeuLeuLysSerGlnGlnLeuLeuThrAlaSerGlnArgAsn 346
QY 1548 TTAAGCTTCAGAGATTATACGATCTCGTATGATGAGCTTAACTACTTCAACAATCTTC 1607
Db 347 LeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeu 366
QY 1608 GATGCTTTTGGTATTATGCACTATACCTTAACTGAAAAGTAGAGATTAATCAAGATGAC 1667
Db 367 AspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlnAspAsnHisAspLys 386
QY 1668 ACCAACCGCATATACCGTTTATATGCGCAAGCGAAGCGAAGAGAGAAATGCTAGTAC 1727
Db 387 AsnAsnAlaGlyValAlaThrValTyrMetGlyLysArgProGlnGlnLysAsnAlaSerTyr 406
QY 1728 CATTAAGCT 1736
Db 407 HisLeuAla 409

```

```

RX PubMed=14679231; DOI=10.1128/JB.186.1.110-121.2004;
RA Kaia A., Beesen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234134; AAF3954.1; -; Genomic_DNA.
DR HSSP; Q53284; 1C4P.
DR SMR; Q7X0Y3; 63-398.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaeinnogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KM kinase
SQ SEQUENCE 440 AA; 50342 MW; B3C6DBA50DB2C39 CRC64;

Alignment Scores:
Pred. No.: 1e-113 Length: 440
Score: 1723.00 Matches: 327
Percent Similarity: 92.43% Conservative: 27
Best Local Similarity: 85.38% Mismatches: 29
Query Match: 46.24% Indels: 0
DB: Gaps: 2

US-09-940-235-12 (1-2096) x Q7X0Y3_STRPY (1-440)
QY 588 ATTGCTGACCTGAGTGGCTGTGACCGTGCATCTGTCAACCAAGCGAATGGTGTG 647
Db 27 IleAlaGlyTyrGlnTyrPheLeuAspArgProSerValAsnAsnSerGlnLeuVal 46
QY 648 AGCGTTGCTGTACTGTGGAGGGAGCAATCAAGCATTAATCTTAATTTTGAATC 707
Db 47 SerMetAlaGlyIleValGlnGlyThrAspLysValPheIleAsnPheGlnIle 66
QY 708 GATCTAACATCAACGACCGCTCATGAGAGAAAGACAGGCAAGGCTTAAGTCAAAAATCA 767
Db 67 AspLeuThrSerGlnProAlaHisGlyGlyLysThrGlnGlnGlnLysLeuSerProLysSer 86
QY 768 AAACATTTGCTACTGATAGTGGCGCATGTCATATAAATTGAGAAAGCTGACTCTA 827
Db 87 LysProPheAlaThrAspAsnGlyAlaMetProHisLysLeuGlnLysAlaAspLeuLeu 106
QY 828 AAGCTTATTGAAGAACATTTGATGCTTACGTCACAGTAAAGCACTACTTTGAGGTC 887
Db 107 LysAlaIleGlnGlnArgLeuIleAlaAsnValHisSerAsnAspGlyTyrPheGlnVal 126
QY 888 ATTGATTTTGCAAGCGATGCAACCATTAAGTGAAGCGCAAGCTCTACTTGTGAC 947
Db 127 IleAspPheAlaSerAspAlaThrIleThrAspArgAspAsnIleTyrPheAlaAsn 146
QY 948 AAGATGTTGCGTAACTTTCGCCGACCACTGTCCAGAAATTTTGTGAAGCGGACAT 1007
Db 147 GlnAsglySerValThrLeuProThrGlnProIleGlnGlnPheLeuAlaGlyHis 166
QY 1008 GTGGCCGTTAGAACATTAAGAAAAACCAATCAAAACCAAGCGAAATCTGTGATGTG 1067
Db 167 ValArgValAlaGlyProTyrGlySgIlySgIlyProIleGlnThrProAlaLysSerValAspIle 186
QY 1068 GAATATACGTGTACAGTTACTCCCTTAACCCGTAGACGATTTCAGACAGGCTGAAA 1127
Db 187 ArgTyrThrValGlnPheThrProLeuAsnProAspAspPheLysProValLeuLys 206
QY 1128 GATATTAAGCTATTGAAAAACATAGCTATCGGTGACACATCAATCTCAAGATTACTA 1187
Db 207 AspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeu 226
QY 1188 GCTCAAGCACAAGCAATTTTAAACAAAACCAACCCAGGCTATACGATTATGAACGTGAC 1247
Db 227 AlaGlnAlaGlnSerIleLeuAsnGlnSerHisSerAspTyrThrIleTyrGlnArgAsp 246
QY 1248 TCCTCAATGCTACATGACATGACATTTTTCGTAAGCTTTCACCAATGATGACAGAG 1307

```


Qy	1608	GATGCTTTGGTATTATGACTATACCTTA	CTGGAAGTAGAGATAATCAGATGAC	1667
Db	367	AspAlaPheaspIle	metAspTyrThrLeuThrGlySerValGluAspAsnHisAspLys	386
Qy	1668	ACCAACCGTATCATACCGTTTATATGGGGAAGCGACCCGAAGAGAGAATGCTAGCTAC	1727	
Db	387	AsnAsnArgValValThrValTyrMetGlyLysArgProLysGlyAlaLysGlySerTyr	406	
Qy	1728	CATTAGCT	1736	
Db	407	HisLeuAla	409	

Search completed: January 28, 2006, 02:34:30
 Job time : 245.882 secs